

GenCore version 5.1.9
 Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2006, 01:58:45 ; Search time 14.2605 Seconds
 (without alignments)
 1315.251 Million cell updates/sec

Title: US-10-612-318-2
 Perfect score: 4313
 Sequence: 1 MEEVGRWGLKPLRLAGL.....SSHPALHTTSDSAGVQTEF 821

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 99297 seqs, 22845552 residues

Total number of hits satisfying chosen parameters: 99297

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published Applications AA New:

- 1: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US09_NEW_PUB_PUB.pap:*
- 2: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US06_NEW_PUB_PUB.pap:*
- 3: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US07_NEW_PUB_PUB.pap:*
- 4: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US08_NEW_PUB_PUB.pap:*
- 5: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US09_PCT_NEW_PUB_PUB.pap:*
- 6: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US10_NEW_PUB_PUB.pap:*
- 7: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US11_NEW_PUB_PUB.pap:*
- 8: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US60_NEW_PUB_PUB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

lt	o.	Score	Query Match	Length	DB ID	Description
1	658.5	15.3	383	7	US-11-293-697-3206	Sequence 3206, Ap
2	522.5	19.4	620	7	US-11-293-697-4145	Sequence 4145, Ap
3	225	5.2	547	6	US-10-449-902-39758	Sequence 39758, A
4	158	3.7	1949	7	US-11-289-102-317	Sequence 317, App
5	154	3.6	675	6	US-10-449-902-48447	Sequence 48447, A
6	154	3.6	675	6	US-10-449-902-50618	Sequence 50618, A
7	145	3.4	1259	7	US-11-223-945-40	Sequence 40, Appl
8	142.5	3.3	210	6	US-10-953-349-23295	Sequence 23295, A
9	138.5	3.2	471	7	US-11-293-697-2588	Sequence 2588, Ap
10	136.5	3.2	3113	6	US-10-505-928-325	Sequence 325, App
11	135.5	3.1	911	6	US-10-449-902-50530	Sequence 50530, A
12	134	3.1	190	6	US-10-953-349-15678	Sequence 15678, A
13	134	3.1	210	6	US-10-953-349-15676	Sequence 15676, A
14	134	3.1	849	6	US-10-953-349-1432	Sequence 1432, Ap
15	134	3.1	1003	6	US-10-953-349-1431	Sequence 1431, Ap
16	134	3.1	1392	6	US-10-953-349-1430	Sequence 1430, Ap
17	133.5	3.1	1234	7	US-11-327-900-6	Sequence 6, Appl
18	133.5	3.1	1384	6	US-10-505-928-799	Sequence 799, App
19	130.5	3.0	788	6	US-10-449-902-41167	Sequence 41167, A
20	128	3.0	240	6	US-10-953-349-18175	Sequence 18175, A
21	128	3.0	246	6	US-10-953-349-18174	Sequence 18174, A
22	127	2.9	698	6	US-10-449-902-53993	Sequence 53993, A
23	126.5	2.9	681	6	US-10-953-349-10919	Sequence 10919, A
24	126.5	2.9	1328	6	US-10-504-973-32	Sequence 32, Appl
25	126.5	2.9	1570	7	US-11-312-958-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1

US-11-293-697-3206
 ; Sequence 3206, Application US/11293697
 ; Publication No. US20060105376A1
 ; GENERAL INFORMATION:

; APPLICANT: HELIX RESEARCH INSTITUTE
 ; TITLE OF INVENTION: Novel full length cDNA
 ; FILE REFERENCE: HI-A0106
 ; CURRENT APPLICATION NUMBER: US/11/293.697
 ; CURRENT FILING DATE: 2005-12-05
 ; PRIOR APPLICATION NUMBER: US/10/108.260
 ; PRIOR FILING DATE: 2002-03-28
 ; NUMBER OF SEQ ID NOS: 5458
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3206
 ; LENGTH: 383
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-11-293-697-3206

Query Match 15.3%; Score 658.5; DB 7; Length 383;

Best Local Similarity 35.4%; Pred. NO. 3.6e-32;
 Matches 146; Conservative 81; Mismatches 145; Indels 41; Gaps 4;

Qy	280	EGPAPDAIRMYRQSKELYGTWELCGNEVOILSNLMBELGPELKAELGRLKGPQER	339
Db	7	EAQAFLEAVOFFRQEKHGHGSMITGDEIQLL	39
Qy	340	QRWQIQISDAVHYMYEAKARFEVLKSVQVQAPQAVIRTDMDQIITKEHLAKIR	399
Db	40	-----EAYTLVQHVSGLSKALKECRALTGLGRTIRSDMDQIVNSKNYLKIKK	91
Qy	400	AFILPKAEVCRVNHVOPITPSTILEALMVPTSCQFTFVRDVFPEVTDMLNVLNVEGGDK	459
Db	92	AWVAQPAEKSCLESVOPFVLASILEELMGVSSGFSEVRLFKENVSQNFQTTKVSQ	151
Qy	460	LGEYMEKLSRLAVHPLKMSQCYEKMESI--RLDGLQORFDVSSTSVFKQRAQIHRQMD	517
Db	152	LKEHLDRMLNPLHVSVMPECYTKVNLHLERLDLKSFRFFHIDLVVQRTONYMOELME	211
Qy	518	NAVYTFETLHQBELGKPTKEELCKSORULERVLYKDYDSSSVKRFPEALLOISIP	577
Db	212	NAVTFEQLLSPHLOGEASKTAV--AIEKVKLRVQDYDSDSTIRKKIFPEALVQITLP	269
Qy	578	FLKKLAPTCKSELPRFOELIFEDFARFILVENTYEVVLQTMKDILQAVKEAAVORKH	637
Db	270	TVQKALASTCKPELOKYOQIFADHTNMHVENVYIEHLQILLODETLLKVKIEAAILKKH	329

Wed Jun 28 10:20:22 2006

QY 638 NLYRDSVMNPNHLLAEGAPIDWCEYNSGGGSPSPSTPESATLSEK 690
Db 330 NLFEDNMLPSE--SVSSITDLKPTTGNQASPARASAILFCVILGSETLSNE 380

RESULT 2
US-11-293-697-4145
; Sequence 4145, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HL-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4145
LENGTH: 620
TYPE: PRT
ORGANISM: Homo sapiens
; 1-293-697-4145

Query Match 12.1%; Score 522.5; DB 7; Length 620;
Best Local Similarity 24.7%; Pred. No. 8.5e-24;
Matches 156; Conservative 108; Mismatches 318; Indels 49; Gaps 6;

QY 99 ILTEFLQFYEDQYGVALFNMRHEIEGTGLPQALLMRKVPDLERIVFSGNLFPQHQEDSK 158
Db 24 LLRNFLPCYRQLAASVLRQISREL-GPOETGSQLRSKKLPVREHRRGPIQLQRHPP 82

QY 159 KVRNRFSLPHNYGLVLYENKAAVERQVPPRAVINSAGYKILTSVDOYLELIGNSLP--- 215
Db 83 RMQPIFCVLRGDDGLEWFSKKEEYNGHCGSLGTALTYLTLSQREYLRLLDLCALCESL 142

QY 216 GTTAKSGSAPILKCTOPFLILMHPYARHYFFCMWTEAEQDKQWAVLODCIRHCNGIPE 275
Db 143 GQHTQEEPSLLEVPVSPFLFLQHPFRHLCSAATREAHAWRLALOGGIRLOQTVLOR 202

QY 276 DSKVEGPAFTAIRMYRSKELYGTWEMLCGNEVQILSNLMEELGPELKAELGPKLKGK 335
Db 203 SQAPARAFLDAVRLYRQHQHFGDDVTLGSDAEVLTAVLMREQLPALRAQTLPLGLGA 262

QY 336 PQERQRMQIOTSDAVYHMVYEQAARFEVLKQVQVQAPMAQAVIRTDMDQIITSKEHLA 395
Db 263 GRARAWATELLDAVHAVALAGASAGLCAFOPEKDELLASLEKTIKPDVDQLLRQARVA 322

QY 396 SKTRAFILPKAEVCVRNHHVQYIPSIIEALMVPTSQGTEVRDVFKEVTDNMLNVINEG 455
Db 323 GLRUTDIRGPLESCLRREVDLPQPRVVQTLRTVBASLEAVRTLLAQGMDRLSHRLROSP 382

QY 456 GIDKLGEYMEKLSRLAYHPLKMQSCYKEMESL--RLDGLAQRFVDSVTSVFQRAQIHRM 513
Db 383 SGTLRREYVSGENPMDLALMQTCTREASRGRGLGQLAAPGFLGMSQSLVFGAQDLAQ 442

QY 514 EQMDNAVYFTTLLHQELGKPTKEBLCKSIORVLRLVKKYDYDSSVKRFPFREALLO 573
Db 443 QLMADAVATFLQADQCTTALNCDAQARLERVRGRVLKKFKSDSLAQRRFIRGWGLC 502

QY 574 ISIPFLKLLKATPKCSLPRFQELIFEDFAFILLVENTYEEVVLQTMKDIILQAVKEAAV 633
Db 503 IFLPFLVLSQLEPGCKELPEFEGDVLAAGSQALTTEGIEDYINGCLLQRIQIDELKLTIG 562

QY 634 QRKHNLYRSMWMSNDNLHLLAEGAPIDW---GBEYSNSGGGSPSPSTPESATLSE 689
Db 563 ANDVSCITLDSCL-----EVPWEQEGADETEAREEGGACPRQPDSCA--- 604

QY 690 KRRRAKQVSVVQDEEVLGPEAFESPSPPPA 720
Db 605 -----QIQPLCPPPS 614

RESULT 3
US-10-449-902-39758
; Sequence 39758, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39758
LENGTH: 547
TYPE: PRT
ORGANISM: Oryza sativa
; US-10-449-902-39758

Query Match 5.2%; Score 225; DB 6; Length 547;
Best Local Similarity 21.1%; Pred. No. 3.2e-06;
Matches 121; Conservative 105; Mismatches 235; Indels 112; Gaps 24;

QY 98 KILTEFLQFYEDQYGVALFN--SMRHEIEGTGLPQALLMRKVPDLERIVFSGNLFP 151
Db 20 KLEKFAKNVSMYLYVAMKASIKQVKAAGKWSLL--KCPKEPDAPDRK---TGHML 74

QY 152 QHQEDSKKRNRFSLPHNYGLVLYENKAAVERQVPPRAVINSAGYKILTSV-DOYLELI 210
Db 75 KRSKHLKMKNRFFVVRGNLVDYMETEQQKAGKPRGTLNLSGMRIRDPNDITYKKI 134

QY 211 GNSLPGTTAKSGSAPILKCTOPFLILMHPYARHYFFCMWTEAEQDK--WQAVLODCIRHC 269
Db 135 LDLAACKNMLNVDLPKPEVLPDFSLEIHHB--TREIVLLOCENADDDHKSMDLDDC-RWY 192

QY 270 NNGIP-EDSKVEGPAFTAIRMYRSKELYGTWEMLCGNEVQI---LSNLWHEELGPELK 325
Db 193 SGELNISDDKIHQIIFQWALWRSRQSGINGYSSGGGGEQYITDCINEAIVDAVMPESVD 252

QY 326 AELGPRKLGKQPER-----QRMQIOTSDAVYHMVYEQAARFEVLKQV 370
Db 253 AKLTMPWAVRSKIRNTFTSNVNTFVSTGVKPAWAAAYGV--KKVRDEVKAKIAEGVAPIA 311

QY 371 QVOPAMQAVIRTDMDQIITSKEHLASKIRAFILPKAEVCVRNHHVQYIPSIIEALMVPTS 430
Db 312 DAQASVQAKILEWVDK--PSKDAIGEK-----VAPHLNPLLDIIIFAPVN 353

QY 431 QGFTVEVRDVPFKEVT-----DMNLNVINEGGIDKLGEYMEKLSRLAYHPL-KMOS 479
Db 354 EAFKIVLSAYDTALTGRKERYTTREDRMHYMNMHANNSEFPWADRLKMDL-FDPLWAMRT 412

QY 480 CYEKMESLRDLGQQRFDVSSVTFKQRAQIHRMEQMDNAVYFTTLLHQELGK----- 533
Db 413 VFSDSVPMGVTG-----KARRRLKTKLNNALFTFETRLSEAGGCKHMDRI 457

QY 534 -GPTKEEL---CK-STORVLRLV---KKYDYDSSVKRFPFREALLOISIFLKLKLAIP 585
Db 458 AAETRETLLKADCRASIGRVLGKILFGVVEVFWDTLIIRPSRALVGLADAIIDAIPKQVD 517

QY 586 TKSELPRFQELIFEDFAFILLVENTYEEVVLQ 618
Db 518 P-----ED-----ILERTLNNILIQ 532

RESULT 4
US-11-289-102-317

Wed Jun 28 10:20:22 2006

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM protein - protein search, using sw model
Run on: June 24, 2006, 01:56:20 ; Search time 38.7072 Seconds
(without alignments)
1856.571 Million cell updates/sec

Title: US-10-612-318-2
Perfect score: 4313
Sequence: 1 MEEVGRWGLKPLRLGLRG.....SSHPLHTTTSDSAGVQTF 821

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patent's AA.*

- 1: /EMC_Celerra_SIDS3/ptodata/2/iaa/5_COMB.pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/iaa/6_COMB.pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/iaa/7_COMB.pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/iaa/H_COMB.pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/iaa/pCTUS_COMB.pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/iaa/RE_COMB.pep.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	176.5	4.1	2090	US-09-538-092-1081	Sequence 1081, Ap
2	176.5	4.1	2120	US-09-949-016-9768	Sequence 9768, Ap
3	159.5	3.7	1248	US-09-080-897-2	Sequence 2, Appli
4	159.5	3.7	1248	US-09-323-735-2	Sequence 2, Appli
5	158.5	3.7	1315	US-08-899-595-3	Sequence 3, Appli
6	152	3.5	864	US-09-538-092-1269	Sequence 1268, Ap
7	151	3.5	1255	US-09-080-897-4	Sequence 4, Appli
8	151	3.5	1255	US-08-899-595-1	Sequence 1, Appli
9	151	3.5	1255	US-09-323-735-4	Sequence 4, Appli
10	147	3.4	688	US-09-141-047-8	Sequence 8, Appli
11	147	3.4	1274	US-09-095-443-2	Sequence 2, Appli
12	147	3.4	1636	US-10-433-794-1	Sequence 1, Appli
13	146.5	3.4	693	US-09-081-385-154	Sequence 154, App
14	146.5	3.4	693	US-09-752-639-154	Sequence 154, App
15	146.5	3.4	693	US-09-712-813-154	Sequence 154, App
16	146.5	3.4	693	US-09-700-354A-154	Sequence 154, App
17	144	3.3	983	US-09-311-236-2	Sequence 2, Appli
18	144	3.3	983	US-09-912-917-2	Sequence 2, Appli
19	144	3.3	1444	US-09-902-540-16727	Sequence 16727, A
20	143.5	3.3	916	US-09-949-016-11417	Sequence 11417, A
21	142.5	3.3	915	US-09-538-092-863	Sequence 863, App
22	142.5	3.3	916	US-09-949-016-6611	Sequence 6611, Ap
23	142.5	3.3	1020	US-09-538-092-911	Sequence 911, App
24	139.5	3.2	638	US-09-949-016-10708	Sequence 10708, A
25	137.5	3.2	912	PCT-US95-03747-2	Sequence 2, Appli
26	137.5	3.2	1231	US-08-714-741-41	Sequence 41, Appli

27	136.5	3.2	2482	1	US-08-328-254-6	Sequence 6, Appli
28	136.5	3.2	3210	2	US-09-538-092-1154	Sequence 1154, Ap
29	136	3.2	603	2	US-09-906-779-4	Sequence 4, Appli
30	136	3.2	694	2	US-09-949-016-7117	Sequence 7117, Ap
31	134	3.1	641	2	US-08-961-083-160	Sequence 160, App
32	134	3.1	641	2	US-09-536-784-160	Sequence 160, App
33	134	3.1	641	2	US-09-765-271-160	Sequence 160, App
34	134	3.1	641	2	US-09-765-272A-160	Sequence 160, App
35	134	3.1	863	2	US-10-067-457-5	Sequence 5, Appli
36	133	3.1	868	2	US-09-800-729-106	Sequence 106, App
37	133	3.1	921	2	US-09-800-729-199	Sequence 199, App
38	132.5	3.1	1388	1	US-08-685-576-1	Sequence 1, Appli
39	132	3.1	802	2	US-09-823-240A-2	Sequence 2, Appli
40	132	3.1	1312	2	US-09-041-886-19	Sequence 19, Appli
41	132	3.1	1312	2	US-09-648-281-2	Sequence 19, Appli
42	132	3.1	1312	2	US-09-707-919A-19	Sequence 3, Appli
43	132	3.1	1312	2	US-09-083-268-3	Sequence 3, Appli
44	132	3.1	1312	2	US-08-981-998A-3	Sequence 3, Appli
45	131.5	3.0	885	2	US-09-914-259-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1

US-09-538-092-1081
; Sequence 1081, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurafPatSeqFormatter Version 0.9
; SEQ ID NO 1081
; LENGTH: 2090
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number P35658
US-09-538-092-1081

Query Match 4.1%; Score 176.5; DB 2; Length 2090;
Best Local Similarity 19.6%; Pred. No. 5.2e-05;
Matches 148; Conservative 126; Mismatches 258; Indels 225; Gaps 38;
QY 184 RQVPPRAVINSAGYKI-----LTSVDQVLELIGNSLP---GTTAKSGSAPILKCP 230
Db 604 QSAPMSPFSSAKPAASGPIASHPTPLSAPPSSVPLKSSVLPSPGSAQSSSPVPSMW 663
QY 231 TQFPPLIL-----WHPYARHYFFCMTEAEQDKWQ-----AVLODCIRHNCNIGP-DS 277
Db 664 QKSPRITPPAAKPGSPQAKSLQ-PAVAEKQHQWQKSDPVMAGIGEEIAHFQKELELKA 722
QY 278 KVEGPAP-----TDATRWYRSK-----ELYGTWMLCGNEVOILSNLWMBELGPELK 325
Db 723 RTSKACFQVGTSEEMKMLTESDDLHTFLLEIKETTESLHGDLSLKTLLLEGAGVEEA 782
QY 326 AELGPRLLKGPQROQWQIQISDAVY-HMVVE-----QAKARFEVLSKVQVQOPANQAV 379
Db 783 RE-----QNERNR-----DSGLHLILYKPLDPKSEALQIRLRLHYVVFVQV-- 826
QY 380 IRTDMDQIITSK--EHLASKIRAFILPKAEVNCVNVHVPYIPISILEALMWPTSGQFTEVR 437

Db 827 ---DVNDVLDLWDQHLQKXK-----QRR-----LLVPE-----R 854

Qy 438 DVFFKEVTDMLNVLNINEGGIDKLGWEYMEKLSRLAYHPLKMQSCYERKESLRDLGLQORFD 497

Db 855 ETUFLNLAN-NREIIN-----QQRKLNHL-----VDSLQQLRL 887

Qy 498 VSTSVFKORAOIHMRQMDNAVYTFETLLHQELGKGTPKEELCKSIQRLVRLVKKYDY 557

Db 888 YKOTSLWSLSSAV---PSQSSIHSPDSDL-----ESLCNAL---LKTIESHTK 930

Qy 558 DSSSVKRP--FRALLOISIPFLKLAAPTCKSELP-----RFOELIFEDFAR 604

Db 931 SLPKVPKLSLPMKQALR---NFLAKRKTTPVRSTAPASLSRSAFLSQRYEDL-DEVSS 986

Qy 605 FILVENTYEEVLTQVMKD---ILOAVKEAAVQRKHNL-----YRDSNMVHNSDP 651

Db 987 TSSVSQSLSESDARTSCDKDEAVVQAPRHAPVVRTPSIQPSLLPHAAFPKSHLVHGSSP 1046

Qy 652 NL-----HLLAEGAPIDWGEYSNSGGGSPSPSTPESA-----685

Db 1047 GVMGTSVATSSASKIIPQGA---DSTMLATKTVKHGAPSPHPSHPISAPQQLAAALRRQMASQ 1104

Qy 686 ---TLSEKRRRAKQVSVVQDEEV-----GLPFEASPESSPPASPDGVTE 727

Db 1105 APAVNTLTS---TLKNVPQVNVQELKNPAPTPSTAMGSSVPYSTAKTPHPVLTTPVAANO 1162

Qy 728 IR-GLLAQGLRPESPPA-GPLINGAPAGESPOPKAA---PEASSPPASPLQHLPLGKA 781

Db 1163 AKQGLSLNLSKPSGTPASGQSLSGDKASGTAKIETAVTSTPSASQGFSPFSFGTG 1222

Qy 782 VDLG--PPKPSDQETGEQVSSPSHPALHTTTEDSAG 816

Db 1223 FNGLIITPSSNFTAAQATPSTKSSQPDASFSSG 1259

RESULT 2

US-09-949-016-9768

; Sequence 9768, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-03

; PRIOR FILING DATE: 2000-10-03

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 9768

; LENGTH: 2120

; TYPE: PRT

; ORGANISM: Human

US-09-949-016-9768

Query Match 4.1%; Score 176.5; DB 2; Length 2120;

Best local similarity 19.6%; Pred. No. 5.4e-05;

Matches 148; Conservative 126; Mismatches 258; Indels 225; Gaps 38;

Qy 184 ROVPPRAVINGAGYKI-----LTSVDQYLELIGNSLP---GTTAKSGSAPILKCP 230

Db 634 QSAPPNPPSSASKAPASGSLSHPTPLSAPPSSVPLKSSVLPSPSGRSAGSSSPVPSMV 693

Qy 231 TQPLIL-----WHPYARHYFCMTAEQDKWQ-----AVLQDCIRHCNNGIPE-DS 277

Db 694 QKSPRITPPAKPGSPQAKSLQ-PAVAEKQGHQWSDPVMAGIEIAHFOKELELKA 752

Qy 278 KVBGPAP----TDAIRMYRQSK-----ELYGTWMLCGNEVQILSNLWMLGELPELK 325

Db 753 RTSKACFOVGTSEMKMLRTESDHLHTFLLEIKETTESLHGDISSLKTLLLEGFAGVEEA 812

Qy 326 AELGRLKGPQERQOWTOISDANY-HMVE-----QAKARPEEVLKSVQVQVPMQAV 379

Db 813 RE-----QNERNR-----DSGYLHLLYRRPLDPSEAOQEIRLRLHGVKFAVQ-- 856

Qy 380 IRITMDQIITGK--EHLAKIRAFILPKAEVVCVRNHVQYPIPSILEALMVPSTQSGFTVR 437

Db 857 ---DVNDVLDLWDQHLQKXK-----QRR-----LLVPE-----R 884

Qy 438 DVFFKEVTDMLNVLNINEGGIDKLGWEYMEKLSRLAYHPLKMQSCYERKESLRDLGLQORFD 497

Db 885 ETUFLNLAN-NREIIN-----QQRKLNHL-----VDSLQQLRL 917

Qy 498 VSTSVFKORAOIHMRQMDNAVYTFETLLHQELGKGTPKEELCKSIQRLVRLVKKYDY 557

Db 918 YKOTSLWSLSSAV---PSQSSIHSPDSDL-----ESLCNAL---LKTIESHTK 960

Qy 558 DSSSVKRP--FRALLOISIPFLKLAAPTCKSELP-----RFOELIFEDFAR 604

Db 961 SLPKVPKLSLPMKQALR---NFLAKRKTTPVRSTAPASLSRSAFLSQRYEDL-DEVSS 1016

Qy 605 FILVENTYEEVLTQVMKD---ILOAVKEAAVQRKHNL-----YRDSNMVHNSDP 651

Db 1017 TSSVSQSLSESDARTSCDKDEAVVQAPRHAPVVRTPSIQPSLLPHAAFPKSHLVHGSSP 1076

Qy 652 NL-----HLLAEGAPIDWGEYSNSGGGSPSPSTPESA-----685

Db 1077 GVMGTSVATSSASKIIPQGA---DSTMLATKTVKHGAPSPHPSHPISAPQQLAAALRRQMASQ 1134

Qy 686 ---TLSEKRRRAKQVSVVQDEEV-----GLPFEASPESSPPASPDGVTE 727

Db 1135 APAVNTLTS---TLKNVPQVNVQELKNPAPTPSTAMGSSVPYSTAKTPHPVLTTPVAANO 1192

Qy 728 IR-GLLAQGLRPESPPA-GPLINGAPAGESPOPKAA---PEASSPPASPLQHLPLGKA 781

Db 1193 AKQGLSLNLSKPSGTPASGQSLSGDKASGTAKIETAVTSTPSASQGFSPFSFGTG 1252

Qy 782 VDLG--PPKPSDQETGEQVSSPSHPALHTTTEDSAG 816

Db 1253 FNGLIITPSSNFTAAQATPSTKSSQPDASFSSG 1289

RESULT 3

US-09-080-897-2

; Sequence 2, Application US/09080897

; Patent No. 5985574

; GENERAL INFORMATION:

; APPLICANT: King, Mary-Claire

; APPLICANT: Lynch, Eric D.

; APPLICANT: Lee, Ming

; APPLICANT: Morrow, Jan E.

; APPLICANT: Welch, Piri L.

; APPLICANT: Leon, Pedro E.

; TITLE OF INVENTION: Modulators of Actin

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 75 DENISE DRIVE

; CITY: HILLSBOROUGH

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94010

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; FILING DATE: US/09/080,897

; CLASSIFICATION:

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM protein - protein search, using sw model
Run on: June 24, 2006, 01:45:00 ; Search time 189.971 Seconds
(without alignments)
3997.655 Million cell updates/sec
Title: US-10-612-318-2
Perfect score: 4313
Sequence: 1 MEEVGRWGLKPLRGLRG.....SSHPALHTTDSAGVQTEF 821

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 2849598 seqs, 925015592 residues
Total number of hits satisfying chosen parameters: 2849598
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : Uniprot 7.2.*
1: uniprot_sprot.*
2: uniprot_trembl.*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	3799	88.1	733	1 NIBL HUMAN	Q96tal homo sapien
2	3799	88.1	733	2 Q5VWV6 HUMAN	Q5VWV6 homo sapien
3	3791	87.9	739	2 Q4LE55 HUMAN	Q4LE55 homo sapien
4	3791	87.9	746	2 Q5VWV7 HUMAN	Q5VWV7 homo sapien
5	3432.5	79.6	749	1 NIBL MOUSE	Q8rlf1 mus musculu
6	3432.5	79.6	749	2 Q543S7 MOUSE	Q543s7 m 4 days ne
7	3426.5	79.4	749	2 Q8BQ71 MOUSE	Q8Bq71 mus musculu
8	3426.5	79.4	749	2 Q3U813 MOUSE	Q3Uc84 mus musculu
9	3423.5	79.4	749	2 Q3UCV6 MOUSE	Q3Uc84 mus musculu
10	3420.5	79.3	749	2 Q3UCV6 MOUSE	Q3Uc84 mus musculu
11	3420.5	79.3	749	2 Q3U814 MOUSE	Q3Uc84 mus musculu
12	3416.5	79.2	749	2 Q3UDW4 MOUSE	Q3Ud4 mus musculu
13	3310.5	76.8	706	2 Q8CC78 MOUSE	Q8cc78 mus musculu
14	2589	60.0	526	2 Q6NRY7 HUMAN	Q6nry7 homo sapien
15	2182.5	50.6	888	2 Q6GPK6 XENLA	Q6gpk6 xenopus lae
16	2027	47.0	396	2 Q9H8K1 HUMAN	Q9h8k1 homo sapien
17	1943	45.0	631	2 Q4T3H2 TETNG	Q4t3h2 tetraodon n
18	1761.5	40.8	392	2 Q3MHT9 RAT	Q3mht9 rattus norv
19	1617.5	37.5	700	2 Q4REAL TETNG	Q4real tetraodon n
20	1425	33.0	327	2 Q2YD88 HUMAN	Q2y888 homo sapien
21	1166	27.0	928	2 Q2TTR2 HUMAN	Q2ttr2 homo sapien
22	1164	27.0	928	1 NIBA HUMAN	Q9bzq8 homo sapien
23	1161	26.9	926	2 Q3UW53 MOUSE	Q3uw53 mus musculu
24	1085.5	25.2	845	2 Q6P879 MOUSE	Q6pe79 mus musculu
25	1048	24.3	201	2 Q9H6L6 HUMAN	Q9h6l6 homo sapien
26	1044	24.2	795	2 Q5F374 CHICK	Q5f374 gallus gall
27	955.5	22.2	584	2 Q3TDH5 MOUSE	Q3tdh5 mus musculu
28	916.5	21.7	277	2 Q4T8L1 TETNG	Q4t8l1 tetraodon n
29	848	19.7	690	2 Q5RG84 BRARE	Q5rg84 brachydanio
30	827.5	19.2	725	2 Q9ESL7 MOUSE	Q9esl7 mus musculu
31	818.5	19.0	736	2 Q9ESN0 RAT	Q9esn0 rattus norv

32 755 17.5 592 2 Q4S226 TETNG
33 530.5 12.3 666 2 Q86XR4 HUMAN
34 522.5 12.1 620 2 Q8N894 HUMAN
35 519.5 12.0 697 2 Q86XR2 HUMAN
36 518.5 12.0 651 2 Q86XR3 HUMAN
37 468 10.9 579 2 Q8TEQ3 HUMAN
38 430.5 10.0 591 2 Q2NL48 RAT
39 421.5 9.8 596 2 Q7Z6H6 HUMAN
40 415.5 9.6 510 2 Q3TD68 MOUSE
41 312.5 7.2 400 2 Q3TQH6 MOUSE
42 256 5.9 528 2 Q5SG75 DICDI
43 220 5.1 99 2 Q5TEV9 HUMAN
44 183 4.2 665 2 Q4Q387 LEIMA
45 178.5 4.1 2080 2 Q3KQZ0 HUMAN

ALIGNMENTS

RESULT 1
NIBL HUMAN
ID NIBL HUMAN STANDARD; PRT; 733 AA.
AC Q96TAL; Q9BUS2; Q9NT35;
DT 13-AUG-2002, integrated into UniProtKB/Swiss-Prot.
DT 13-AUG-2002, sequence version 2.
DT 07-FEB-2006, entry version 28.
DE Niban-like protein (Meg-3).
GN Name=C9orf88;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RT Miyata T., Inagi R., Yasuda Y., Kurokawa K.;
RA "Homo sapiens meg-3 mRNA, complete cds."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 38-733.
RC The German CDNA consortium;
RG Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 278-733.
RC TISSUE=Cervix;
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hoptkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP PHOSPHORYLATION SITES SER-652; SER-679 AND SER-683.
RA PubMed=15302935; DOI=10.1073/pnas.0404720101;
RA Beausoleil S.A., Jedrychowski M., Schwartz D., Elias J.E., Villen J.,
Li J., Cohn M.A., Cantley L.C., Gygi S.P.;
RA "Large-scale characterization of HeLa cell nuclear phosphoproteins";

RL	Proc. Natl. Acad. Sci. U.S.A. 101:12130-12135(2004).	
CC	-/- SIMILARITY: Belongs to the Niban family.	
CC	-/- SIMILARITY: Contains 1 PH domain.	
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms	
CC	Distributed under the Creative Commons Attribution-NoDerivs License	
CC	-----	
DR	EMBL; AF151783; AAK57556.1; -; mRNA.	
DR	EMBL; AL137555; CAB70809.1; -; mRNA.	
DR	EMBL; BC001979; AAB01979.2; -; mRNA.	
DR	PIR; T46394.	
DR	Ensembl; ENSG00000136830; Homo sapiens.	
DR	HGNC; HGNC:25282; C9orf88.	
DR	InterPro; IPR001849; PH.	
DR	InterPro; IPR011993; PH type.	
DR	PROSITE; PS00003; PH_DOMAIN; 1.	
KW	Phosphorylation.	
FT	CHAIN 1 733 Niban-like protein.	
FT	/FTID=PRO_0000213121.	
FT	DOMAIN 55 179 PH.	
FT	MOD_RES 552 552 Phosphoserine.	
FT	MOD_RES 679 679 Phosphoserine.	
FT	MOD_RES 683 683 Phosphoserine.	
FT	CONFLICT 305 305 H -> L (in Ref. 1).	
FT	CONFLICT 664 664 P -> L (in Ref. 3).	
FT	SEQUENCE 733 AA; 82683 MW; 38E1C24CF737F3DB CRC64;	
Query Match	88.1%; Score 3799; DB 1; Length 733;	
Best Local Similarity	96.1%; Pred. No. 7.5e-190;	
Matches 733; Conservative	0; Mismatches 0; Indels 30; Gaps 1;	
QY	59 MGNMGGORSTSLWPWTPGVSEFLEGGMLWFQOEKTKILTEFLQFYEDQYGVALLFNS 118	
DB	1 MGNMG-----EXTGKILTEFLQFYEDQYGVALLFNS 30	
QY	119 MRHEIEGTGLPQALLWRKVPDLDERIVFSGNLFQHQEDSKKWRNRFSLVPHNYGLVLYEN 178	
DB	31 MRHEIEGTGLPQALLWRKVPDLDERIVFSGNLFQHQEDSKKWRNRFSLVPHNYGLVLYEN 90	
QY	179 KAAERQVPPRAVINSAGYKILTSVDQVLELIGNSLPGTTAKSGSAPILKCTPFPILW 238	
DB	91 KAAERQVPPRAVINSAGYKILTSVDQVLELIGNSLPGTTAKSGSAPILKCTPFPILW 150	
QY	239 HPYARHYFCMTEAEQDKQWAVLQDCIRHCNNGIPEDSKVGEPAFTDAIRMYRSKELY 298	
DB	151 HPYARHYFCMTEAEQDKQWAVLQDCIRHCNNGIPEDSKVGEPAFTDAIRMYRSKELY 210	
QY	299 GTWEMLCGNEVOILSNLWMEELGPALKELGPRKQKQORQOWIQTSDAVYHMVYEQ 358	
DB	211 GTWEMLCGNEVOILSNLWMEELGPALKELGPRKQKQORQOWIQTSDAVYHMVYEQ 270	
QY	359 KARFEVLISKVQVQVQAMQAVIRTDMDQIITSKEHLASKIRAFILPKAEVCVRNHVQYI 418	
DB	271 KARFEVLISKVQVQVQAMQAVIRTDMDQIITSKEHLASKIRAFILPKAEVCVRNHVQYI 330	
QY	419 PSTLEALAVPTSGQTEVRDVFKEVTDNMLVINEGGIDKLGEYMEKLSRLAYHPLMQ 478	
DB	331 PSTLEALAVPTSGQTEVRDVFKEVTDNMLVINEGGIDKLGEYMEKLSRLAYHPLMQ 390	
QY	479 SCYEKMSRLDGLQORFDVSSVFKQRAQIHMREQMDNAVYTFETLLHSELGKGPTKE 538	
DB	391 SCYEKMSRLDGLQORFDVSSVFKQRAQIHMREQMDNAVYTFETLLHSELGKGPTKE 450	
QY	539 ELCKSTQRLVRLVKKYDYDSSSVKRRKPFREALQISIPFLKKLAPTKCSLPRFOELI 598	
DB	451 ELCKSTQRLVRLVKKYDYDSSSVKRRKPFREALQISIPFLKKLAPTKCSLPRFOELI 510	
QY	599 FEDFARFILVENTVEEVVLOTVMKDILQAVKEAAVQRKHNLYRDSWMNSDPNHLILAE 658	
DB	511 FEDFARFILVENTVEEVVLOTVMKDILQAVKEAAVQRKHNLYRDSWMNSDPNHLILAE 570	
QY	659 GABIDWGEYSNSGGGSPSPSTPESATLSEKRRAKQVSVVQDEVGLPPEASPEPPP 718	

Db	571 GABIDWGEYSNSGGGSPSPSTPESATLSEKRRAKQVSVVQDEVGLPPEASPEPPP 630
QY	719 PASPDGVTEIRGLLAQGLRPESPPAGPLINGAPAGESPQKAAPEASSPPASPLQHLPL 778
Db	631 PASPDGVTEIRGLLAQGLRPESPPAGPLINGAPAGESPQKAAPEASSPPASPLQHLPL 690
QY	779 GKAVDGLGPKPKSDQBTGEQVSSPSHPALHTTTEDSAGVQTEF 821
Db	691 GKAVDGLGPKPKSDQBTGEQVSSPSHPALHTTTEDSAGVQTEF 733

RESULT 2

Q5VVW6 HUMAN	PRELIMINARY; PRT; 733 AA.
Q5VVW6;	
AC	Q5VVW6;
DT	10-MAY-2005, integrated into UniProtKB/TREMBL.
DT	10-MAY-2005, sequence version 1.
DT	07-FEB-2006, entry version 6.
DE	OTTHUM00000022187.
GN	Name=RP11-356B19.6; ORFNames=RP11-356B19.6-003;
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC	Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RA	Laird G.;
RL	Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN	[2]
RP	NUCLEOTIDE SEQUENCE.
RA	Babbage A.;
RL	Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC	Distributed under the Creative Commons Attribution-NoDerivs License
CC	-----
EMBL	AL445222; CAH72915.1; -; Genomic DNA.
EMBL	AL390116; CAH72915.1; JOINED; Genomic DNA.
EMBL	AL390116; CAH73085.1; -; Genomic DNA.
EMBL	AL445222; CAH73085.1; JOINED; Genomic DNA.
Ensembl	ENSG00000136830; Homo sapiens.
InterPro	IPR001849; PH.
InterPro	IPR011993; PH type.
PROSITE	PS00003; PH_DOMAIN; 1.
SEQUENCE	733 AA; 82683 MW; 38E1C24CF737F3DB CRC64;

Query Match 88.1%; Score 3799; DB 2; Length 733;

Best Local Similarity 96.1%; Pred. No. 7.5e-190;

Matches 733; Conservative 0; Mismatches 0; Indels 30; Gaps 1;

QY	59 MGNMGGORSTSLWPWTPGVSEFLEGGMLWFQOEKTKILTEFLQFYEDQYGVALLFNS 118
DB	1 MGNMG-----EXTGKILTEFLQFYEDQYGVALLFNS 30
QY	119 MRHEIEGTGLPQALLWRKVPDLDERIVFSGNLFQHQEDSKKWRNRFSLVPHNYGLVLYEN 178
DB	31 MRHEIEGTGLPQALLWRKVPDLDERIVFSGNLFQHQEDSKKWRNRFSLVPHNYGLVLYEN 90
QY	179 KAAERQVPPRAVINSAGYKILTSVDQVLELIGNSLPGTTAKSGSAPILKCTPFPILW 238
DB	91 KAAERQVPPRAVINSAGYKILTSVDQVLELIGNSLPGTTAKSGSAPILKCTPFPILW 150
QY	239 HPYARHYFCMTEAEQDKQWAVLQDCIRHCNNGIPEDSKVGEPAFTDAIRMYRSKELY 298
DB	151 HPYARHYFCMTEAEQDKQWAVLQDCIRHCNNGIPEDSKVGEPAFTDAIRMYRSKELY 210
QY	299 GTWEMLCGNEVOILSNLWMEELGPALKELGPRKQKQORQOWIQTSDAVYHMVYEQ 358
DB	211 GTWEMLCGNEVOILSNLWMEELGPALKELGPRKQKQORQOWIQTSDAVYHMVYEQ 270
QY	359 KARFEVLISKVQVQVQAMQAVIRTDMDQIITSKEHLASKIRAFILPKAEVCVRNHVQYI 418
DB	271 KARFEVLISKVQVQVQAMQAVIRTDMDQIITSKEHLASKIRAFILPKAEVCVRNHVQYI 330
QY	419 PSTLEALAVPTSGQTEVRDVFKEVTDNMLVINEGGIDKLGEYMEKLSRLAYHPLMQ 478
DB	331 PSTLEALAVPTSGQTEVRDVFKEVTDNMLVINEGGIDKLGEYMEKLSRLAYHPLMQ 390
QY	479 SCYEKMSRLDGLQORFDVSSVFKQRAQIHMREQMDNAVYTFETLLHSELGKGPTKE 538
DB	391 SCYEKMSRLDGLQORFDVSSVFKQRAQIHMREQMDNAVYTFETLLHSELGKGPTKE 450
QY	539 ELCKSTQRLVRLVKKYDYDSSSVKRRKPFREALQISIPFLKKLAPTKCSLPRFOELI 598
DB	451 ELCKSTQRLVRLVKKYDYDSSSVKRRKPFREALQISIPFLKKLAPTKCSLPRFOELI 510
QY	599 FEDFARFILVENTVEEVVLOTVMKDILQAVKEAAVQRKHNLYRDSWMNSDPNHLILAE 658
DB	511 FEDFARFILVENTVEEVVLOTVMKDILQAVKEAAVQRKHNLYRDSWMNSDPNHLILAE 570
QY	659 GABIDWGEYSNSGGGSPSPSTPESATLSEKRRAKQVSVVQDEVGLPPEASPEPPP 718

GenCore version 5.1.9
Copyright (C) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2006, 01:50:05 ; Search time 31.0676 Seconds
(without alignments)
2542.646 Million cell updates/sec

Title: US-10-612-318-2
Perfect score: 4313
Sequence: 1 MEEVGGRWGLKPLRLAGL.....SSHPALHTTSDSAGVQTF 821

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : p12_80
1: p12_80
2: p12_80
3: p12_80
4: p12_80

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3624	84.0	696	2 T46394	hypothetical prote
2	176.5	4.1	2090	2 S26058	probable transform
3	162	3.8	858	2 S15762	neurofilament trip
4	161.5	3.7	1259	2 T16038	hypothetical prote
5	156	3.6	1461	2 T41643	probable involveme
6	152.5	3.5	1087	1 QP4SH	neurofilament trip
7	151	3.5	864	2 A40671	dynamain, internal
8	151	3.5	1255	2 T31065	diaphanous protein
9	148	3.4	1017	2 PC4035	cell-cycle-depende
10	147	3.4	990	2 T14756	hypothetical prote
11	147	3.4	7962	2 I38346	elastic titin - hu
12	146.5	3.4	1400	2 T31555	hypothetical prote
13	145.5	3.4	802	2 S44382	PM1 protein, splic
14	145.5	3.4	1130	2 T34081	hypothetical prote
15	145	3.4	1018	2 T43168	hypothetical prote
16	145	3.4	1072	1 A37221	neurofilament trip
17	145	3.4	1260	1 TVRTNU	protein-tyrosine k
18	144.5	3.4	871	2 T48502	hypothetical prote
19	144	3.3	798	2 I50479	neurofilament medi
20	143.5	3.3	1251	2 A56677	neuronal cell cycl
21	142.5	3.3	802	2 S42518	PM1 protein, splic
22	142.5	3.3	851	2 B40671	dynamain, internal
23	142.5	3.3	916	2 Q27864	neurofilament trip
24	142.5	3.3	1020	1 QFHUH	neurofilament trip
25	141	3.3	581	2 T22341	hypothetical prote
26	139	3.2	2259	2 S29236	calcium channel pr
27	138.5	3.2	268	2 H84684	En/Spm-like transp
28	138.5	3.2	1282	2 JE0120	Glycoprotein A - m
29	138.5	3.2	1388	2 S74245	serine/threonine-s

ALIGNMENTS

RESULT 1

T46394

hypothetical protein DKFZp434H0820.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004

R:Ottenwälder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23031

A:Accession: T46394

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-696 <AAA>

A:CROSS-references: UNIPROT:Q96TA1; UNIPARC:UPI000016ACC1; EMBL:AL137555

A:Experimental source: adult testis; clone DKFZp434H0820

C:Genetics:

A>Note: DKFZp434H0820.1

Query Match 84.0%; Score 3624; DB 2; Length 696;

Best Local Similarity 100.0%; Pred. No. 1.6e-199;

Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	126	TGLFQAQLLRKVPILDERIVFSGNLFOHEDSKWKWRNFSLVPHNYGLVLYENKAAAYEQ	185
DB	1	TGLFQAQLLRKVPILDERIVFSGNLFOHEDSKWKWRNFSLVPHNYGLVLYENKAAAYEQ	60
QY	186	VPPRAVINSAGYKILTSVDVLELIGNSLPQTAKSGSAPILKCTPTOPPLILMHPYARHY	245
DB	61	VPPRAVINSAGYKILTSVDVLELIGNSLPQTAKSGSAPILKCTPTOPPLILMHPYARHY	120
QY	246	YFCMWTAEQDKQWQVLDQICIRHCNNGIPEDSKVEGPAFTDAIRMYRSKELYGTWEMLC	305
DB	121	YFCMWTAEQDKQWQVLDQICIRHCNNGIPEDSKVEGPAFTDAIRMYRSKELYGTWEMLC	180
QY	306	GNEVQILSNLMEEELGPELKAELGPRKLGKQPOROWIOISDAVYHVMYEQAKARPEEV	365
DB	181	GNEVQILSNLMEEELGPELKAELGPRKLGKQPOROWIOISDAVYHVMYEQAKARPEEV	240
QY	366	LSKVQVQVQPMQAVITDMDQIITSKEHLASKIRAFILPKAEVCVRNVHQPYPISILEAL	425
DB	241	LSKVQVQVQPMQAVITDMDQIITSKEHLASKIRAFILPKAEVCVRNVHQPYPISILEAL	300
QY	426	MYPTSQGTEVRDVFKEVTDNMLNVINEGGIDKLGEYMEKLSRLAYHPLKMQSCYEKME	485
DB	301	MYPTSQGTEVRDVFKEVTDNMLNVINEGGIDKLGEYMEKLSRLAYHPLKMQSCYEKME	360
QY	486	SRLDGLQORFDVSSVSVEKQRAQIHMEQMDNAVITFTLLHQLGKGPTEBELCKSIQ	545
DB	361	SRLDGLQORFDVSSVSVEKQRAQIHMEQMDNAVITFTLLHQLGKGPTEBELCKSIQ	420
QY	546	RYLRLVLLKYDYDSSSVRRFRFPFALLQISIPFLKLLAPTCKSELPFQELIFEDPARF	605

Db 421 RVLRLVKLYDYSSVRKRRFFREALQISIPFLKLLKLAAPTCKSELPRFOELIFEDFARF 480
Qy 606 ILVENTYEEVLTQVMKDILOAVKEAAVORKHNLVYRDSVMHNSDPNHLHLAGAPIDWG 665
Db 481 ILVENTYEEVLTQVMKDILOAVKEAAVORKHNLVYRDSVMHNSDPNHLHLAGAPIDWG 540
Qy 666 EYNSNGGGSPSPSPESATLSKRRRAKQVSVVQDEVEGLPPEASPEPPSPDG 725
Db 541 EYNSNGGGSPSPSPESATLSKRRRAKQVSVVQDEVEGLPPEASPEPPSPDG 600
Qy 726 TEIRGLLAQGLRPESPPAGPLNGAPAGESPOKPAEPASSPPAPLOHLLPGKAVDLG 785
Db 601 TEIRGLLAQGLRPESPPAGPLNGAPAGESPOKPAEPASSPPAPLOHLLPGKAVDLG 660
Qy 786 PKPDSQDTGTVSSPSHPALHTTTEDSAGVTEF 821
Db 661 PKPDSQDTGTVSSPSHPALHTTTEDSAGVTEF 696

RESULT 2
S26058
probable transforming protein (can) - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S26058
R: von Lindern, M.; Fornetod, M.; van Baal, S.; Jaegle, M.; de Wit, T.; Buijs, A.; Grosve
Mol. Cell. Biol. 12, 1687-1697, 1992
A:Title: The translocation (6;9), associated with a specific subtype of acute myeloid le
k-can mRNA.
A:Reference number: S26058; MUID:92195315; PMID:1549122
A:Accession: S26058
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2090 <VON>
A:Cross-references: UNIPROT:P35658; UNIPARC:UPI000012FC1E; EMBL:X64228; NID:q29652; PIDN

Query Match 4.1%; Score 176.5; DB 2; Length 2090;
Best Local Similarity 19.6%; Pred. NO. 0.11;
Matches 148; Conservative 126; Mismatches 258; Indels 225; Gaps 38;
Qy 184 RQVPPRAVNSGKYI-----LTSVDQYLELIGNSLP---GTTAKSGAPILKCP 230
Db 604 QSAPPSPSSASKPAASGSLHPTLSAPSSVPLKSLVLPSPGRSAGSSSPVSMV 663
Qy 231 TOPPLTL-----NHPVARYHYFCMMTEAQDKWQ-----AVLQDCIRHCNNGIPE-DS 277
Db 664 QKSPRITPPAKPGSPQAKSLQ-PAVAEKQGHQWSDPVVMAGIGEBIAHFQKLEELKA 722
Qy 278 KVEGPAF---TDAIRMYRSK-----ELYGTWMLCGNEVQILSNLWMLGPELK 325
Db 723 RTSKACFQVGTSEEMKMLRTESDDLHTFLEIKETTESLHGDISSLTKTLLEGFAGVEEA 782
Qy 326 AELGPRLKGPOROROWIQISDAVY-HMVYE-----QAKARFEVLSKVQVQVQAMQAV 379
Db 783 RE-----QNERN-----DSGYLHLLYKRPDLDPKSEAQLQIRLHLHVQKFAVQ-- 826
Qy 380 IRTDMDQIITSK--EHLASKTRAFILPKAEVVRHVQPYIPSTILELMAVPTSQGTEVR 437
Db 827 ---DVNDVLDLSDWQHLBQKKK-----QRH-----LIVPE-----R 854
Qy 438 DVPEKEVTDMLNLVNEGIDKLGVEYMEKLSRLAVHPLKMQSCYKRMESLRDGLQRF 497
Db 855 ETUFTNLN-NREIN-----QQRKLNHL-----VDSLQQLRL 887
Qy 498 VSTSVFKORAIQHMREQMDNAVYTFETLHQLBELGKGTKEELCKSORVLERVLKKYDY 557
Db 888 YKQTSLSLSSAV---PSQSIHSFSDL-----ESLCLNLA---LKTTHESHK 930
Qy 558 DSSSVKRP--FREALLQISIPFLKLLKLAAPTCKSEL-----RFOELIFEDFAR 604
Db 931 SLPKVPKAKUSPMKQAKLR---NFLAKRKTTPVNSTAPASLSRSFSLQRYEDL-DEVSS 986
Qy 605 FILVENTYEEVLTQVMKD---ILQAVKEAAVORKHNL-----YRDSVMHNSDP 651

Db 987 TSSVSOSLESSEDARTCKDEAVVQAPRHPVVRTPTSIQPSLLPHAAPAKSHLVHGSSP 1046
Qy 652 NL-----HLLAEGAPIDMGEYNSNGGGSPSPSPESA----- 685
Db 1047 GVNGTSTVATSKIIPOGA--DSTMLATKTVKHGAPSPSPHPIISAPOQLAAALRRQMASQ 1104
Qy 686 ----TLSEKRRRAKQVSVVQDEVE-----GLPFEASPEPPSPDGVTE 727
Db 1105 APAVNTLTS--TLKNPQVNVVQELKNNPATPSTAMGSSVPYSTAKTPHPVLTVAANO 1162
Qy 728 IR-GLLAQGLRPESPPA-GLLNGAPAGESQPKAA-----PEASSPPAPLOHLLPGKA 781
Db 1163 AKQGLSLNLSKPGTTPASGQLSGDKASGTAKIETAVTSTPSSQGFSPSPSGTG 1222
Qy 782 VDLG--PPKPSDOETGEOVSSPSHPALHTTTEDSAG 816
Db 1223 FNFGIITPTSSNFTTAOQATPSTKSSQDPAFSSGG 1259

RESULT 3
S15762
neurofilament triplet M protein - chicken
C:Species: Gallus gallus (chicken)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S15762; S08061; A27040
R: Zopf, D.; Dineva, B.; Betz, H.; Gundelfinger, E.D.
Nucleic Acids Res. 18, 521-529, 1990
A:Title: Isolation of the chicken middle-molecular weight neurofilament (NF-M) gene and
A:Reference number: S15762; MUID:90174973; PMID:2106668
A:Accession: S15762
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-858 <ZOP>
A:Cross-references: UNIPROT:P16053; UNIPARC:UPI000017137C; EMBL:X17102; NID:g63688; PID
R: Zopf, D.; Dineva, B.; Betz, H.; Gundelfinger, E.
submitted to the EMBL Data Library, November 1989
A:Reference number: S08061
A:Accession: S08061
A:Molecule type: DNA
A:Residues: 1-355,368-858 <ZOP>
A:Cross-references: UNIPARC:UPI00001774FC; EMBL:X17102
R: Zopf, D.; Hermans-Borgmeyer, I.; Gundelfinger, E.D.; Betz, H.
Genes Dev. 1, 699-708, 1987
A:Title: Identification of gene products expressed in the developing chick visual syste
A:Reference number: A27040; MUID:88112814; PMID:3123320
A:Accession: A27040
A:Molecule type: mRNA
A:Residues: 350-546, 'R', 548-858 <ZOP>
A:Cross-references: UNIPARC:UPI000017137B; GB:X05558; NID:g63685; PIDN:CAA29073.1; PID:

Query Match 3.8%; Score 162; DB 2; Length 858;
Best Local Similarity 21.1%; Pred. NO. 0.19;
Matches 136; Conservative 92; Mismatches 260; Indels 156; Gaps 29;
Qy 293 QSKELYGTWMLC--GNEVOILSNL-----VMEEGLPELKAELGPRILKGPQE 338
Db 79 QSSLNGAAELKLSRNEKEQLOGLNDRFAGYIEKHVYLEQONKEIEALA-ALRQKHAG 137
Qy 339 ROROWIQISDAVHYVMYEQAKARFEVLSKVQVQ----- 373
Db 138 R-----AQLGDA-YEQELRELRLGALQVSHKAKAQTOLDSEHIEDIQRERFEDEALRD 192
Qy 374 --PAMQAVIRPDMDQIITSKEHLASKIRAFILPKAEVVCVRNVHVPYIPSIIEALMVPTSQ 431
Db 193 ETEATIAARKMEEAASLRAELDKKVS--LQDEVAFLAGNHEEVEAELLAQLQ--ASH 248
Qy 432 GTEVRDVF-----FKEV-----TDMNLNVINGE---GDKLGEYMEK----- 466

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2006, 01:44:35 ; Search time 145.152 Seconds
(without alignments)
2586.080 Million cell updates/sec

Title: US-10-612-318-2

Perfect score: 4313

Sequence: 1 MEEVGRWGLKPLRLAGLRG.....SSHPALHTTTDSAGVQTEF 821

Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

- 1: Genesep21980s:**
- 2: Genesep21990s:**
- 3: Genesep22000s:**
- 4: Genesep22001s:**
- 5: Genesep22002s:**
- 6: Genesep22003as:**
- 7: Genesep22003bs:**
- 8: Genesep22004s:**
- 9: Genesep22005s:**
- 10: Genesep22006s:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4313	100.0	821	ADL23466	Adl23466 Human STR
2	4114	95.4	791	ADL23468	Adl23468 Human STR
3	3799	88.1	733	ADOS5194	Ados5194 Protein #
4	3791	87.9	746	AAMS1209	Aams1209 Human ost
5	3791	87.9	746	ABM81075	Abm81075 Tumour-as
6	3788	87.8	733	AAB11456	Aab11456 Human Meg
7	3662	84.9	703	ABP62933	Abp62933 Human pol
8	2027	47.0	396	AAB95443	Aab95443 Human pro
9	1164	27.0	928	AAB67451	Aab67451 Amino aci
10	1164	27.0	928	AAB94034	Aab94034 Human pro
11	1164	27.0	928	ADY17313	Ady17313 PRO polyp
12	1164	27.0	928	ADY17311	Ady17311 PRO polyp
13	1164	27.0	928	AEE04903	Aee04903 Cancer-as
14	1164	27.0	928	AEE04901	Aee04901 Cancer-as
15	1164	27.0	928	AEG07465	Aeg07465 Human aco
16	1160	26.9	928	ADS10972	Ads10972 Human the
17	989.5	22.9	726	AAM84036	Aam84036 Human imm
18	838	19.4	726	ADB82699	Adb82699 Human pro
19	827.5	19.2	725	AEE04898	Aee04898 Cancer-as
20	658.5	15.3	383	ADM04521	Adm04521 Human pro
21	658.5	15.3	383	REC87451	Rec87451 Human CRN
22	595	13.8	298	AB42724	Ab42724 Human ORF
23	542.5	12.6	703	ABR83120	Ab83120 Human BCN

24	522.5	12.1	620	7	ADM05460	Adm05460 Human pro
25	522.5	12.1	620	9	AEC88390	Aec88390 Human BCN
26	519.5	12.0	697	7	ABR83122	Ab83122 Human BCN
27	518.5	12.0	651	7	ABR83121	Ab83121 Human BCN
28	468	10.9	579	7	ADD67630	Add67630 Human Lyl
29	468	10.9	579	7	ABR83123	Ab83123 Human BCN
30	329	7.6	247	4	AAM90354	Aam90354 Human imm
31	244	5.7	177	2	AAV21849	Aav21849 Human sig
32	177	4.1	2130	9	ADX07781	Adx07781 Cyclin-de
33	176.5	4.1	2090	10	AEF58534	Aef58534 Human nuc
34	176.5	4.1	2127	7	ADF09542	Adf09542 Human pol
35	176.5	4.1	2058	4	AAB97070	Aab97070 Human pol
36	175.5	4.1	2058	4	AAB97070	Aab97070 Human pol
37	166	3.8	31	9	ADV13344	Adv13344 Human pho
38	162.5	3.8	873	3	ADO55126	Ado55126 Protein #
39	161.5	3.7	1259	8	ADN23114	Adn23114 Bacterial
40	159.5	3.7	1082	7	ADE14368	Ade14368 Human int
41	159.5	3.7	1088	8	ADH61284	Adh61284 INTSIG pr
42	159.5	3.7	1248	2	AAV13464	Aav13464 Human dia
43	158.5	3.7	1112	8	ADR66113	Adr66113 Human pro
44	158.5	3.7	1112	8	ADR66455	Adr66455 Human pro
45	158.5	3.7	1112	9	ADY18523	Ady18523 PRO polyp

ALIGNMENTS

RESULT 1

ADL23466
ID ADL23466 standard; protein; 821 AA.

AC ADL23466;

DT 01-JUL-2004 (first entry)

XX Human STR50 long variant polypeptide.

XX Human; STR50; long variant; neurotoxic stress; neurodegenerative disease;
stroke; Parkinson's disease; Alzheimer's disease; Huntington's disease;
tumour; cancer; cerebroprotective; vasotropic; antiparkinsonian;
neuroprotective; nootropic; anticonvulsant.

OS Homo sapiens.

XX US2004068098-A1.

XX 08-APR-2004.

XX 01-JUL-2003; 2003US-00612318.

XX 01-JUL-2002; 2002US-0393251P.

XX (KACH/) KACHALSKY S G.

XX (FAER/) FAERMAN A.

XX (PELO/) PEL-OR Y.

XX Kachalsky SG, Faerman A, Pel-Or Y;

XX WPI; 2004-304622/28.

XX N-PSDB; ADL23465.

XX New STR50 polynucleotides and polypeptides, useful in diagnosing and
treating neurodegenerative diseases, e.g. stroke, Parkinson's disease,
Alzheimer's disease or Huntington's disease and tumours.

XX Claim 14; SEQ ID NO 2; 38pp; English.

XX The invention relates to a purified polynucleotide encoding the human
STR50 polypeptide. The purified polypeptide has the biological activity
to modulate neurotoxic stress. The polynucleotide and the polypeptide and
compositions of the invention are useful in diagnosing and treating
neurodegenerative diseases, e.g. stroke, Parkinson's disease, Alzheimer's
disease, Huntington's disease and tumours. This sequence represents the

CC	human STR50 long variant polypeptide of the invention.	
XX	Sequence 821 AA;	
SQ	Query Match 100.0%; Score 4313; DB 8; Length 821;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 821; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MEEVGRWGLKPRLAGLGTGSDCPALAAQQOTALKDVILGVRNWASSGPFLLGLSLWRMG 60	DE
DB	1 MEEVGRWGLKPRLAGLGTGSDCPALAAQQOTALKDVILGVRNWASSGPFLLGLSLWRMG 60	XX
QY	61 WNGGQSTSLWPWTGPFVSEFLEDDGGMGLWFOEQKTKILTEFLQFYEDQYGVAFNSMR 120	KW
DB	61 WNGGQSTSLWPWTGPFVSEFLEDDGGMGLWFOEQKTKILTEFLQFYEDQYGVAFNSMR 120	KW
QY	121 HEIETGLPQALLWRKVPDLDERIVFSGNLFQHQEDSKKWRNRFSLVPHNYGLVLYENKA 180	KW
DB	121 HEIETGLPQALLWRKVPDLDERIVFSGNLFQHQEDSKKWRNRFSLVPHNYGLVLYENKA 180	KW
QY	181 AYERQVPPRAVINSAGYKILTSVDQYLEIGNSLPGTTAKSGSAPILKCPQFPLILWHP 240	KW
DB	181 AYERQVPPRAVINSAGYKILTSVDQYLEIGNSLPGTTAKSGSAPILKCPQFPLILWHP 240	KW
QY	241 YARHYFFCMWTEAEQDKQWAVLQDCIRHCNNGIPEDSKVEGPAFTDAIRMYRQSKELYGT 300	OS
DB	241 YARHYFFCMWTEAEQDKQWAVLQDCIRHCNNGIPEDSKVEGPAFTDAIRMYRQSKELYGT 300	XX
QY	301 WMLCGNEVQILSNLWMEELGPELKAELGPRKQKPOERQWQIQLSDAVHYMYEQAKA 360	XX
DB	301 WMLCGNEVQILSNLWMEELGPELKAELGPRKQKPOERQWQIQLSDAVHYMYEQAKA 360	XX
QY	361 RFEVLVKVQVQPAWQAVIRTDMDQIITSKEHLASKIRAFILPKAEVCVRNHVQPIPS 420	XX
DB	361 RFEVLVKVQVQPAWQAVIRTDMDQIITSKEHLASKIRAFILPKAEVCVRNHVQPIPS 420	XX
QY	421 ILEALMVPTSGQFTEVRDVFKEVTDMLNVLNVEGGIDKLGEYMEKLSRLAYHPLKMQSC 480	XX
DB	421 ILEALMVPTSGQFTEVRDVFKEVTDMLNVLNVEGGIDKLGEYMEKLSRLAYHPLKMQSC 480	XX
QY	481 YKMSLSRLDLQORFVSTSVFKORAIHMRBMONAVTTEILLHOELGKGPTEEL 540	XX
DB	481 YKMSLSRLDLQORFVSTSVFKORAIHMRBMONAVTTEILLHOELGKGPTEEL 540	XX
QY	541 CKSIQVLERVLKYYDYDSSSVKRRFFREALQIISIPFLKLLKLAFTCKSELPRFOELIFE 600	XX
DB	541 CKSIQVLERVLKYYDYDSSSVKRRFFREALQIISIPFLKLLKLAFTCKSELPRFOELIFE 600	XX
QY	601 DFARFILVENTYEEVVLQTVMKDILQAVKEAAVQKKNLYRDSWVHNSDPNHLHLLAEGA 660	XX
DB	601 DFARFILVENTYEEVVLQTVMKDILQAVKEAAVQKKNLYRDSWVHNSDPNHLHLLAEGA 660	XX
QY	661 PIDWGEYSNCGGSGSPSPSTPESATLSEKRRRAKQVSVVQDEEVGLPFEASPESSPPA 720	XX
DB	661 PIDWGEYSNCGGSGSPSPSTPESATLSEKRRRAKQVSVVQDEEVGLPFEASPESSPPA 720	XX
QY	721 SPGVTEIRGLLAQGLRPESPPAGPLNGAPAGESPQKAAPEASSPPASPIQLHLPGK 780	XX
DB	721 SPGVTEIRGLLAQGLRPESPPAGPLNGAPAGESPQKAAPEASSPPASPIQLHLPGK 780	XX
QY	781 AVDLGPPKPSDQETGEQVSSPSHPALHTTTEDSAGVOTEF 821	XX
DB	781 AVDLGPPKPSDQETGEQVSSPSHPALHTTTEDSAGVOTEF 821	XX
RESULT 2		
ADL23468		
ID	ADL23468 standard; protein: 791 AA.	
XX		
AC	ADL23468;	
XX		
DT	01-JUL-2004 (first entry)	
XX		

Human STR50 short variant polypeptide.

Human, STR50; short variant; neurotoxic stress; neurodegenerative disease; stroke; Parkinson's disease; Alzheimer's disease; Huntington's disease; tumour; cancer; cerebroprotective; vasotropic; antiparkinsonian; neuroprotective; nootropic; anticonvulsant.

Homo sapiens.

US2004068098-A1.

08-APR-2004.

01-JUL-2003; 2003US-00612318.

01-JUL-2002; 2002US-0393251P.

(KACH/) KACHALSKY S G.

(FAER/) FAERMAN A.

(PELO/) PEL-OR Y.

Kachalsky SG, Faerman A, Pel-Or Y;

WPI; 2004-304622/28.

N-PSDB; ADL23467.

New STR50 polynucleotides and polypeptides, useful in diagnosing and treating neurodegenerative diseases, e.g. stroke, Parkinson's disease, Alzheimer's disease or Huntington's disease and tumours.

Claim 15; SEQ ID NO 4; 38pp; English.

The invention relates to a purified polynucleotide encoding the human STR50 polypeptide. The purified polypeptide has the biological activity to modulate neurotoxic stress. The polynucleotide and the polypeptide and compositions of the invention are useful in diagnosing and treating neurodegenerative diseases, e.g. stroke, Parkinson's disease, Alzheimer's disease, Huntington's disease and tumours. This sequence represents the human STR50 short variant polypeptide of the invention.

Sequence 791 AA;

Query Match 95.4%; Score 4114; DB 8; Length 791;

Best Local Similarity 96.3%; Pred. No. 2e-310;

Matches 791; Conservative 0; Mismatches 0; Indels 30; Gaps 1;

QY 1 MEEVGRWGLKPRLAGLGTGSDCPALAAQQOTALKDVILGVRNWASSGPFLLGLSLWRMG 60

DB 1 MEEVGRWGLKPRLAGLGTGSDCPALAAQQOTALKDVILGVRNWASSGPFLLGLSLWRMG 60

QY 61 WNGGQSTSLWPWTGPFVSEFLEDDGGMGLWFOEQKTKILTEFLQFYEDQYGVAFNSMR 120

DB 61 WNGGQSTSLWPWTGPFVSEFLEDDGGMGLWFOEQKTKILTEFLQFYEDQYGVAFNSMR 120

QY 121 HEIETGLPQALLWRKVPDLDERIVFSGNLFQHQEDSKKWRNRFSLVPHNYGLVLYENKA 180

DB 91 HEIETGLPQALLWRKVPDLDERIVFSGNLFQHQEDSKKWRNRFSLVPHNYGLVLYENKA 150

QY 181 AYERQVPPRAVINSAGYKILTSVDQYLEIGNSLPGTTAKSGSAPILKCPQFPLILWHP 240

DB 151 AYERQVPPRAVINSAGYKILTSVDQYLEIGNSLPGTTAKSGSAPILKCPQFPLILWHP 210

QY 241 YARHYFFCMWTEAEQDKQWAVLQDCIRHCNNGIPEDSKVEGPAFTDAIRMYRQSKELYGT 300

DB 211 YARHYFFCMWTEAEQDKQWAVLQDCIRHCNNGIPEDSKVEGPAFTDAIRMYRQSKELYGT 270

QY 301 WMLCGNEVQILSNLWMEELGPELKAELGPRKQKPOERQWQIQLSDAVHYMYEQAKA 360

DB 271 WMLCGNEVQILSNLWMEELGPELKAELGPRKQKPOERQWQIQLSDAVHYMYEQAKA 330

QY 361 RFEVLVKVQVQPAWQAVIRTDMDQIITSKEHLASKIRAFILPKAEVCVRNHVQPIPS 420

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model
Run on: June 24, 2006, 20:00:20 ; Search time 1403 Seconds
(without alignments)
6045.429 Million cell updates/sec

Title: US-10-612-318-3
Perfect score: 4530
Sequence: 1 gggccctgcacaaatgcg.....atttaaaaaaaaaaaaaa 4533

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 93554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA.*
1: /EMC-Celerra_SIDS3/ptodata/2/ina/1 COMB.seq.*
2: /EMC-Celerra_SIDS3/ptodata/2/ina/5 COMB.seq.*
3: /EMC-Celerra_SIDS3/ptodata/2/ina/6A COMB.seq.*
4: /EMC-Celerra_SIDS3/ptodata/2/ina/6B COMB.seq.*
5: /EMC-Celerra_SIDS3/ptodata/2/ina/7 COMB.seq.*
6: /EMC-Celerra_SIDS3/ptodata/2/ina/H COMB.seq.*
7: /EMC-Celerra_SIDS3/ptodata/2/ina/PCTUS COMB.seq.*
8: /EMC-Celerra_SIDS3/ptodata/2/ina/PP COMB.seq.*
9: /EMC-Celerra_SIDS3/ptodata/2/ina/RE COMB.seq.*
10: /EMC-Celerra_SIDS3/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1342.6	29.6	1446	3	US-09-205-258-20 Sequence 20, Appl
2	1342.6	29.6	1446	3	US-10-004-860-20 Sequence 20, Appl
3	648.8	14.3	713	3	US-09-702-705-203 Sequence 203, App
4	648.8	14.3	713	3	US-09-736-457-203 Sequence 203, App
5	648.8	14.3	713	3	US-09-614-124B-203 Sequence 203, App
6	648.8	14.3	713	3	US-09-671-325-203 Sequence 203, App
7	648.8	14.3	713	3	US-09-589-184-203 Sequence 203, App
8	648.8	14.3	713	3	US-09-658-824-203 Sequence 203, App
9	648.8	14.3	713	3	US-10-017-754-203 Sequence 203, App
10	648.8	14.3	713	3	US-09-651-563-203 Sequence 203, App
11	648.8	14.3	713	3	US-09-702-705-202 Sequence 202, App
12	238.4	5.3	283	3	US-09-736-457-202 Sequence 202, App
13	238.4	5.3	283	3	US-09-614-124B-202 Sequence 202, App
14	238.4	5.3	283	3	US-09-671-325-202 Sequence 202, App
15	238.4	5.3	283	3	US-09-589-184-202 Sequence 202, App
16	238.4	5.3	283	3	US-09-658-824-202 Sequence 202, App
17	238.4	5.3	283	3	US-10-017-754-202 Sequence 202, App
18	238.4	5.3	283	3	US-09-651-563-202 Sequence 202, App
19	238.4	5.3	283	3	US-09-519-642-202 Sequence 202, App
20	238.4	5.3	283	3	US-09-519-642-202 Sequence 202, App
21	207	4.6	217	3	US-09-513-999C-10754 Sequence 10754 A
22	155.4	3.4	169	4	US-09-880-107-762 Sequence 762, App
23	85.8	1.9	337	3	US-09-513-999C-33591 Sequence 33591, A

24	66.4	1.5	1588	3	US-09-490-291-7	Sequence 7, Appl
25	63.2	1.4	1995	2	US-08-425-069-3	Sequence 3, Appl
26	63.2	1.4	1995	2	US-08-317-844B-3	Sequence 3, Appl
27	62.6	1.4	2338	2	US-08-425-069-1	Sequence 1, Appl
28	62.6	1.4	2338	2	US-08-317-844B-1	Sequence 1, Appl
29	61.2	1.4	1852	3	US-09-969-852-4	Sequence 4, Appl
30	60.4	1.3	7218	2	US-08-232-463-14	Sequence 14, Appl
31	58	1.3	7994	3	US-09-902-540-797	Sequence 797, App
32	57.4	1.3	1907	4	US-10-094-749-1541	Sequence 1541, Ap
33	57.4	1.3	1914	4	US-10-094-749-1451	Sequence 1451, Ap
34	57.4	1.3	2541	3	US-09-902-540-7900	Sequence 7900, Ap
35	57.4	1.3	6603	3	US-09-902-540-796	Sequence 796, App
36	57.2	1.3	7218	2	US-08-232-463-14	Sequence 14, Appl
37	56.4	1.2	1233	4	US-10-077-381-12	Sequence 12, Appl
38	56.4	1.2	2305	3	US-09-475-515-80	Sequence 80, Appl
39	56.4	1.2	2306	3	US-09-475-515-82	Sequence 82, Appl
40	56.2	1.2	1926	3	US-09-249-585A-2	Sequence 2, Appl
41	56.2	1.2	1926	3	US-09-410-399-3	Sequence 3, Appl
42	56.2	1.2	2580	3	US-09-050-863-2	Sequence 2, Appl
43	56.2	1.2	5452	2	US-09-359-081-2	Sequence 1, Appl
44	56.2	1.2	5452	2	US-09-130-114-1	Sequence 1, Appl
45	56.2	1.2	6644	3	US-08-875-435B-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-09-205-258-20 Application US/09205258
; Sequence 20, Application US/09205258
; Patent No/ 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06

241	Db		GAGAGGAGGTGAAAGCTGTGTGTGAGCCAGCTGGGTTCCGCGCCACTCACCCAGAGCTG	300
3426	Qy		GCTGGGCGAGGACCGGAGAGGAGCACTGCTGCGCCCTCTGGCCCTGCTCTTCGCGAGT	3485
301	Db		GCTGGGCGAGGACCGGAGAGGAGCACTGCTGCGCCCTCTGGCCCTGCTCTTCGCGAGT	360
3486	Qy		TAGGGGTGGACCGAGCCTCGCTTTCCCACTGTTCTTGAGGAGGAGGAGGAGGGGGTC	3545
361	Db		TAGGGGTGGACCGAGCCTCGCTTTCCCACTGTTCTTGAGGAGGAGGAGGAGGGGGTC	420
3546	Qy		TTGAGGCTGGAGCGAGGCTGGGGTGCTGGGTGGAGAGATGAGATTTAGGGGGTGCCCTCA	3605
421	Db		TTGAGGCTGGAGCGAGGCTGGGGTGCTGGGTGGAGAGATGAGATTTAGGGGGTGCCCTCA	480
3606	Qy		TGGGGTGGGCAAGCCTGGGGTGAAATGAGAAAGGCCAGAACTGTCAGGTCTGCGGAGGG	3665
481	Db		TGGGGTGGGCAAGCCTGGGGTGAAATGAGAAAGGCCAGAACTGTCAGGTCTGCGGAGGG	540
3666	Qy		GAAGTGCTGAGTGAAGAGGGGAGCCCC - ATCTCTGGGAGATGCTGGGAGTGAAGTGA	3723
541	Db		GAAGTGCTGAGTGAAGAGGGAGCCCCATCTCTGGGGGATGCTGGGAGTGAAGTGA	600
3724	Qy		AGATGGCTGAGTGAGGGTTATGGGGAGCTGAGGTTTATGGGCCTGTGTATCCCTTCT	3783
601	Db		AGATGGCTGAGTGAGGGTTATGGGGAGCTGAGGTTTATGGGCCTGTGTATCCCTTCT	660
3784	Qy		CCCGGCCGACGCTGCTCCCTGCTGCGCCGCTGCGCCACAGGTCTCCCTCTGGTCCCTG	3843
661	Db		CCCGGCCGACGCTGCTCCCTGCTGCGCCGCTGCGCCACAGGTCTCCCTCTGGTCCCTG	720
3844	Qy		TCCCTCTGGTGTGGGATGAGCGGACGCAAGGGGTGTAAATGGGGCTGGGTTCGTCT	3903
721	Db		TCCCTCTGGTGTGGGATGAGCGGACGCAAGGGGTGTAAATGGGGCTGGGTTCGTCT	780
3904	Qy		TCTAGGCGACCCCGAGTCTCTCAGTGTGTGCTGGGGGCGCGAGCGGGCTCTTGAGG	3963
781	Db		TCTAGGCGACCCCGAGTCTCTCAGTGTGTGCTGGGGGCGCGAGCGGGCTCTTGAGG	840
3964	Qy		GGTACAGTTGGGTGGGCGCCCTCCCTGAGGGTCTGGGGTCAAGCTTTGGCCTCTGCTGCCT	4023
841	Db		GGTACAGTTGGGTGGGCGCCCTCCCTGAGGGTCTGGGGTCAAGCTTTGG - CTCTGCTGCCT	899
4024	Qy		CTCAGTCAACAGTCACTCCCTCTGAAATTCAGTCCCTCTTTGGATGTCTTTGTGAG	4083
900	Db		CTCAGTCAACAGTCACTCCCTCTGAAATTCAGTCCCTCTTTTGGATGTCTTTGTGAG	959
4084	Qy		TCACTCTGGGCTGGCTGCTCCCTCTCAGTCTTCTGTTCTGGGACAGGGTCAAGC	4143
960	Db		TCACTCTGGGCTGGCTGCTCCCTCTCAGTCTTCTGTTCTGGGACAGGGTCAAGC	1019
4144	Qy		CAGATGGGCGCCAGGCTGAGATCCGCCACCCAGAGACCCACAGGCCCTCCCTGTT	4203
1020	Db		CAGATGGGCGCCAGGCTGAGATCCGCCACCCAGAGACCCACAGGCCCTCCCTGTT	1078
4204	Qy		GNTTTCGGGGGCGAGGCGAGAAATGGAATCTTTTGGGTCCCGAGGTGGGGTCCCTC	4263
1079	Db		GCTTTTGGGGGGGCGAGGCGAGAAATGGAATCTTTTGGGTCCCGAGGTGGGGTCCCTC	1138
4264	Qy		CCAGCCTGCACTCTCGCTGCCCTAGACTGCTCCCGAGAGGGGCTGTGACCCACAG	4323
1139	Db		CCAGCCTGCACTCTCGCTGCCCTAGACTGCTCCCGAGAGGGGCTGTGACCCACAG	1198
4324	Qy		GAAATGTGTGGCGCTGCAATACGAGGACCCCGAGCTGCCGAGCCCTGGTTTTTGGCG	4383
1199	Db		GAGCTGTGTGGCGCTGGCACTCAGGAGACCCCGAGCTGCCCGAGCCCTGGTCTCGCG	1258
4384	Qy		CATCTTTTCCCTCTTGTCCCGAAGATTTGGCGCTTTAGTGCCTTTTGGAGGGTTCCTCATC	4443
1259	Db		CATCTCTTCCCTCTTGTCCCGAAGATCTCGCCTCTAGTGCCTTTTGGAGGGTTCCTCATC	1318
4444	Qy		ATCCCTCCCTGATTTGTATTGAAATATTATGACACACTGTTCATGCTTTTACTAATCAA	4503

Query Match	Score	DB 3;	Length	1446;
Best Local Similarity	98.28;	Pred. No. 4e-276;		
Matches 1385;	Conservative	2;	Mismatches 19;	Indels 4; Gaps 3;
QY	3126	CGGGGCGAGGGCTGTGTGGCACCGCCAGGAGCGGGGCCACCTGAGTCACACTTTATTTGGGTT	3185	
DB	1	CGGGGCGAGGGCTGTGTGGCACCGCCAGGAGCGGGGCCACCTGAGTCACACTTTATTTGGGTT	60	
QY	3186	CAGTCAACACTTTCTTCTCCCTGTTTTCTTTCTGTGGGATGATCTCAGATGACGGGC	3245	
DB	61	CAGTCAACACTTTCTTCTCCCTGTTTTCTTTCTGTGGGATGATCTCAGATGACGGGC	120	
QY	3246	TGGTTTTGGGTTTTCTGTGTGTGTCACAGGCTGGACACTGTGGGGGCTGGAAAGCC	3305	
DB	121	TGGTTTTGGGTTTTCTGTGTGTGTCACAGGCTGGACACTGTGGGGGCTGGAAAGCC	180	
QY	3306	CCTCCCTCTCTGTCTCTGTGGCTTCACATCCCTCATGGGTGCTGCATCCTTCTCTGGA	3365	
DB	181	CCTCCCTCTCTGTCTCTGTGGCTTCACATCCCTCATGGGTGCTGCATCCTTCTCTGGA	240	

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 25, 2006, 12:06:30 ; Search time 108 Seconds
(without alignments)
10061.783 Million cell updates/sec

Title: US-10-612-318-3
Perfect score: 4530
Sequence: 1 gggctccctgcacaaatgcg.....atttaaaaaaaaaaaaaa 4533

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 296510 seqs, 119862409 residues

Total number of hits satisfying chosen parameters: 593020

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- Published Applications NA New:
- 1: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US09_NEW_PUB.seq:
 - 2: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US06_NEW_PUB.seq:
 - 3: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US07_NEW_PUB.seq:
 - 4: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US08_NEW_PUB.seq:
 - 5: /EMC_Celerra_SIDS3/ptodata/1/pubpna/PCT_NEW_PUB.seq:
 - 6: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US10_NEW_PUB.seq:
 - 7: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq:
 - 8: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	648.8	14.3	713	7 US-11-301-554-203	Sequence 203, App
2	238.4	5.3	283	7 US-11-301-554-202	Sequence 202, App
3	214.2	4.7	2165	7 US-11-293-697-763	Sequence 763, App
4	175.8	3.9	2073	7 US-11-293-697-1702	Sequence 1702, App
5	57.2	1.3	3236	6 US-10-449-902-12664	Sequence 12664, A
6	56.4	1.2	2219	6 US-10-449-902-11307	Sequence 11307, A
7	56.2	1.2	11978	7 US-11-257-851A-63	Sequence 63, Appl
8	53.4	1.2	1841	6 US-10-511-937-2882	Sequence 2882, App
9	53	1.2	152331	7 US-11-175-714-86	Sequence 86, Appl
10	52.8	1.2	2185	6 US-10-449-902-12280	Sequence 12280, A
11	51.8	1.1	3592	6 US-10-449-902-16770	Sequence 16770, A
12	51	1.1	2796	6 US-10-449-902-19530	Sequence 19530, A
13	50.8	1.1	1275	6 US-10-449-902-11706	Sequence 11706, A
14	50.6	1.1	1766	6 US-10-449-902-10050	Sequence 10050, A
15	50	1.1	1392	6 US-10-449-902-2794	Sequence 2794, App
16	50	1.1	1510	6 US-10-449-902-19240	Sequence 19240, A
17	50	1.1	1575	6 US-10-449-902-12085	Sequence 12085, A
18	49.8	1.1	2479	7 US-11-293-697-1966	Sequence 1966, App
19	49.6	1.1	1185	6 US-10-953-349-29401	Sequence 29401, A
20	49.6	1.1	2953	7 US-11-145-307A-175	Sequence 175, App
21	49.4	1.1	1414	6 US-10-449-902-8473	Sequence 8473, App
22	49.4	1.1	1693	6 US-10-449-902-24457	Sequence 24457, A
23	49.4	1.1	2303	6 US-10-449-902-12529	Sequence 12529, A
24	48.8	1.1	992	6 US-10-449-902-10706	Sequence 10706, A
25	48.8	1.1	1799	7 US-11-293-697-1885	Sequence 1885, App

RESULT 1

US-11-301-554-203
; Sequence 203, Application US/11301554
; Publication No. US20060088527A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McRabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C21
; CURRENT APPLICATION NUMBER: US/11/301.554
; CURRENT FILING DATE: 2005-12-13
; PRIOR APPLICATION NUMBER: US 10/283,017
; PRIOR FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US 10/113,872
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 10/017,754
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 09/902,941
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 09/849,626
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: US 09/736,457
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 09/702,705
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: US 09/677,419
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 09/671,325
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 09/658,824
; PRIOR FILING DATE: 2000-09-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2157
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 203
; LENGTH: 713

ALIGNMENTS

26	48.6	1.1	1108	6	US-10-449-902-25818	Sequence 25818, A
27	48.6	1.1	1135	6	US-10-953-349-34319	Sequence 34319, A
28	48.2	1.1	929	6	US-10-449-902-13000	Sequence 13000, A
29	48.2	1.1	1274	6	US-10-449-902-13651	Sequence 13651, A
30	48.2	1.1	5017	7	US-11-297-660-1	Sequence 1, Appl
31	48	1.1	1079	6	US-10-449-902-1712	Sequence 1712, App
32	48	1.1	1080	6	US-10-449-902-3157	Sequence 3157, App
33	48	1.1	1080	6	US-10-449-902-3787	Sequence 3787, App
34	48	1.1	1671	6	US-10-449-902-2816	Sequence 2816, App
35	48	1.1	1902	6	US-10-449-902-14165	Sequence 14165, A
36	48	1.1	1915	6	US-10-449-902-14465	Sequence 14465, A
37	47.8	1.1	2270	6	US-10-449-902-9653	Sequence 9653, App
38	47.6	1.1	827	6	US-10-449-902-10215	Sequence 10215, A
39	47.6	1.1	1479	6	US-10-953-349-27828	Sequence 27828, A
40	47.4	1.0	1114	6	US-10-449-902-8250	Sequence 8250, App
41	47.4	1.0	1762	6	US-10-449-902-17474	Sequence 17474, A
42	47.4	1.0	2047	6	US-10-449-902-21525	Sequence 21525, A
43	47.4	1.0	2206	6	US-10-449-902-25241	Sequence 25241, A
44	47.4	1.0	2366	6	US-10-449-902-9444	Sequence 9444, App
45	47.4	1.0	3508	7	US-11-071-896-6	Sequence 6, Appl

```

; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C21
; CURRENT FILING DATE: 2005-12-13
; CURRENT APPLICATION NUMBER: US 10/283,017
; PRIOR FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US 10/113,872
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 10/017,754
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 09/902,941
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 09/849,626
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: US 09/736,457
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 09/702,705
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: US 09/677,419
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 09/671,325
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 09/658,824
; PRIOR FILING DATE: 2000-09-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 2157
; SEQ ID NO 202
; LENGTH: 283
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-301-554-202

Query Match      5.3%; Score 238.4; DB 7; Length 283;
Best Local Similarity 95.5%; Pred. No. 6.6e-42;
Matches 256; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY    2470   CTACCGGAGCAGCATGTCTATCAACAACAGCAGCCACCCAACTTGCACCTTGCTGCCGAGGG 2529
Db     275   CTCCTCGGGGAGCATGTCTTCGTTAACAAGTGGCCCCAACCTTGCACCTTGCTGCCGAGGG 216
QY    2530   C-GCCCCCATCGACTGGGGGAGAGGTACAGCAACAGCGCGGGGCGCGCAGCCAGCC 2588
Db     215   CAGCCCCCATCGACTGGGGGAGAGGTACAGCAACAGCGCGGGGCGCGCAGCCAGCC 156
QY    2589   CCAGCACCCCGAGCTCAGCCACCCTCTCGGAAAAGCAGCGCGCGCCGAAGCAGGTGTCT 2648
Db     155   CCAGCACCCCGAGGTACGCCACCCCTCTCGGAAAAGCAGCGCGCGCCGAAGCAGGTGTCT 96
QY    2649   CTGTGTCTCAGGATGAGAGGTGGGGCTGCCCTTTGAGGTAGCCCTGAGTCAACCAC 2708
Db     95     CTGTGTCTCAGGATGAGAGGTGGGGCTGCCCTTTGAGGTAGCCCTGAGTCAACCAC 36
QY    2709   CTGCTGCCCGGAGGTGTCACTGAGAT 2736
Db     35     CTGCTGCCCGGAGGTGTCACTGAGAT 8

RESULT 3
US-11-293-697-763
; Sequence 763, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28

```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 25, 2006, 11:28:14 ; Search time 19662 Seconds
(without alignments)
12891.991 Million cell updates/sec

Title: US-10-612-318-3
Perfect score: 4530
Sequence: 1 gggtccctgcacaaatgcg.....atttaaaaaaaaaaaaaa 4533

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_hcc:*
7: gb_est2:*
8: gb_est7:*
9: gb_est8:*
10: gb_est9:*
11: gb_gsal:*
12: gb_gsal2:*
13: gb_gsal3:*
14: gb_gsal4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3553.4	78.4	3656	6 HSM802294	AL137555 Homo sapi
2	2184.4	48.2	2186	14 AY405558	AY405558 Homo sapi
3	1954.8	43.2	2183	14 AY405559	AY405559 Pan trogl
4	1755.4	38.7	3636	6 AK153489	AK153489 Mus muscu
5	1753.4	38.7	3590	6 AK046583	AK046583 Mus muscu
6	1753.4	38.7	3631	6 AK150642	AK150642 Mus muscu
7	1753.4	38.7	3632	6 AK051396	AK051396 Mus muscu
8	1753.4	38.7	3645	6 AK152207	AK152207 Mus muscu
9	1753.4	38.7	3660	6 AK033735	AK033735 Mus muscu
10	1751.8	38.7	3603	6 AK149888	AK149888 Mus muscu
11	1751.8	38.7	3626	6 AK170512	AK170512 Mus muscu
12	1751.8	38.7	3679	6 AK146953	AK146953 Mus muscu
13	1716.6	37.9	3321	6 AK149818	AK149818 Mus muscu
14	1686.2	37.2	2195	14 AY405560	AY405560 Mus muscu
15	898.2	19.6	985	4 BX383821	BX383821 BX383821
16	854.4	18.9	1063	4 BX362527	BX362527 BX362527
17	833.6	18.4	928	3 BQ939703	BQ939703 AGENCOURT
18	818.8	18.1	954	3 BQ929092	BQ929092 AGENCOURT
19	814	18.0	995	3 BQ945457	BQ945457 AGENCOURT

20	808	17.8	1134	2	BM551984
21	802.8	17.7	981	3	BQ919097
22	802.4	17.7	1118	2	BM552910
23	801.2	17.7	973	3	BQ928844
24	789.4	17.4	1024	2	BG750134
25	785.8	17.3	916	3	BQ679480
26	778.2	17.2	923	3	BQ882950
27	775.8	17.1	922	3	BQ878325
28	770.4	17.0	772	9	CK870576
29	763.8	16.9	1057	3	BQ072980
30	763.6	16.9	854	2	BG333436
31	761	16.8	783	2	BG680041
32	755.8	16.7	871	10	DV817082
33	754.8	16.7	948	3	BU501198
34	753.6	16.6	873	4	BX350373
35	752	16.6	903	5	CD514169
36	749.2	16.5	943	3	BQ893750
37	743.6	16.4	822	2	B1334843
38	743.4	16.4	803	2	BM014205
39	742.8	16.4	928	2	BG680193
40	741.2	16.4	908	3	BU190217
41	731.6	16.2	765	5	CF128461
42	726.8	16.0	771	3	BU688793
43	726.6	16.0	837	7	BE799575
44	724	16.0	758	3	BM969109
45	723.4	16.0	1086	2	BM450097

ALIGNMENTS

RESULT 1
HSM802294 3656 bp mRNA linear HTC 22-SEP-2004
LOCUS Homo sapiens mRNA; cDNA DKFZp434H0820 (from clone DKFZp434H0820).
DEFINITION AL137555
ACCESSION AL137555.1 GI:6808242
VERSION HTCC
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 3656)
AUTHORS Ottenwaelder,B.; Obermaier,B.; Deutschenhaur,S.; Schaidp,A.;
Mewes,H.W.; Weil,B.; Aml,C.; Osanger,A.; Fobo,G.; Han,M. and
Wiemann,S.
CONSRFTM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Medigenomix (Martinsried/Germany) within the CDNA
sequencing consortium of the German Genome Project.
This clone (DKFZp434H0820) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp434H0820
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES

Location/Qualifiers
1..3656
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp434H0820"
/db_xref="taxon:9606"
/clone="DKFZp434H0820"
/tissue type="testis"
/clone_lib="434 (synonym: htes3). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
/note="hypothetical protein, N-terminus truncated"

Wed Jun 28 10:20:23 2006

```

gene
CDS
1. 3656
/gene="DKFZp434H0820"
42. .2092
/gene="DKFZp434H0820"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB70809.1"
/db_xref="GI:6808243"
/db_xref="UniProtKB/Swiss-Prot.O96TAL"
/translation="TGLPQQLLRKVLQPLDIERIVFSNGLPIQHQEDSKKWRNRSLVPH
NYGLVLENKAAVERQVPRAVLSNAGYKLTISVDYQLELINGSPLGTGTTAKSGAPLL
KCPITPLILMHPYARHYFCMWTAEAQDKQAVLQDCIRHCNMTGAEPLSPEDSKVGRQAF
KATIPWQSKELGYTWMELCGNVEQLTSLNLMVEGLPELKAELGRLKGPQRQQRQW
TOISDAVHNVYEQAKPFEVLKQVQYQPMQAVIRTDMDQITISKEHLASKIRAF
LIPAEVCVLRNVHYPITFISLELALWPTFSQGTVEVRDVFPEKFTDMNLNVINSGGIDK
LGYEAKVLSIAVHPLKMQSCIKEMSELRLDGLQQRDVSTSVFKQRAOIHLHREQMD
NAVYITFLHOBGLKGTPEELCKTISKVLRERLKKYIDVSSSVKRFPEALLQIIS
TPILLKKLAPTCKSLPFOELIPFADFAFIVLENTVEEVVITQWVKDILQAVKEAR
QRKHLVRDSVMNHSNDNLHLAEAGPITDWGEYSNSGGGSPSPSTPEATILSEK
RFAKQVSVVDDEVEGLPEASPEPSPDPGVTEIRGLLAGQLRPSPPPPGAPLLN
GAPAGESPQAPAAPEASPPAPSLQHLPLGKAVDLPGLPFKPSDQETGEQVSPSPSHPAL
unappendedCDS"

```

ORIGIN

Query Match	78.4%;	Score 3553.4;	DB 6;	Length 3656;
Best Local Similarity	99.4%;	Pred. No. 0;		
Matches 3586;	Conservative 0;	Mismatches 19;	Indels 3;	Gaps 2;
QY	928	CACGGGCTGCCGACAGCCACAGCTGCTCTGGCGCAAGGTGCCATCTGGACGACGCGATCGT	987	
DB	1	CACGGGCTGCCGACAGCCACAGCTGCTCTGGCGCAAGGTGCCATCTGGACGACGCGATCGT	60	
QY	988	CTTCTCGGGGAACCTCTTCCAGCAACGAGGAGACGAGAACTGCGAGAAACCGCTTCAG	1047	
DB	61	CTTCTCGGGGAACCTCTTCCAGCAACGAGGAGACGAGAACTGCGAGAAACCGCTTCAG	120	
QY	1048	CCTCGTCCGCCCACTACGAGGGCTGGTGCTCTACGAAACAAAGCGGCGCTATGACGGCA	1107	
DB	121	CCTCGTCCGCCCACTACGAGGGCTGGTGCTCTACGAAACAAAGCGGCGCTATGACGGCA	180	
QY	1108	GGTCCCAACGAGCGGCTCATCAACAGTCAGGCTACAAATCTCACTCCGCTGGACCA	1167	
DB	181	GGTCCCAACGAGCGGCTCATCAACAGTCAGGCTACAAATCTCACTCCGCTGGACCA	240	
QY	1168	ATACCTGAGCTCATTTGGCAACTCTCTTACGAGGACCGCAAGGTGCGGCAAGTGCCCC	1227	
DB	241	ATACCTGAGCTCATTTGGCAACTCTTACGAGGACCGCAAGGTGCGGCAAGTGCCCC	300	
QY	1228	CATCTCAAGTGGCCCCACACAGTTCGCGCTCATCTCTGGCATCTTATGCGGCTCACTA	1287	
DB	301	CATCTCAAGTGGCCCCACACAGTTCGCGCTCATCTCTGGCATCTTATGCGGCTCACTA	360	
QY	1288	CTACTTCGATGATGACGAGCGGACGAGCAAGTGGCAGGCTGTGTCGAGGACTG	1347	
DB	361	CTACTTCGATGATGACGAGCGGACGAGCAAGTGGCAGGCTGTGTCGAGGACTG	420	
QY	1348	CATCGGCACTGCAACAATGGAAATCCCTGAGGACTCCAAGGTAGAGGGCCCTCGGTTAC	1407	
DB	421	CATCGGCACTGCAACAATGGAAATCCCTGAGGACTCCAAGGTAGAGGGCCCTCGGTTAC	480	
QY	1408	AGATGCGATCCGATGTACGACAGTCCAAGGAGCTGTACGGCACTCTGGAGATGCTGTG	1467	
DB	481	AGATGCGATCCGATGTACGACAGTCCAAGGAGCTGTACGGCACTCTGGAGATGCTGTG	540	
QY	1468	TGGGAACGAGGTGCAGATCTTGAGCAACTGGTGATGAGGAGCTGGGCCCTGACGTGAA	1527	
DB	541	TGGGAACGAGGTGCAGATCTTGAGCAACTGGTGATGAGGAGCTGGGCCCTGACGTGAA	600	
QY	1528	GGCAGAGCTCCGCCCGCGCTGAAGGGCAACCGCAGGAGCGGACGCGCATGTGGATCCA	1587	
DB	601	GGCAGAGCTCCGCCCGCGCTGAAGGGCAACCGCAGGAGCGGACGCGCATGTGGATCCA	660	
QY	1588	GATCTCGGACGCCGTGTGACACANTGTGTGACGAGGCCAAGGCCGCTTCGAGGAGGT	1647	

661	DB		GATCTCGAGCGCGGTGTAACAATGGTGTACGAGCAGGCCAAGGCGCGCTTCGAGGAGGT	720
1648	QY		GCTGTCCAAGGTGCAGCAGGTGCAGCGCGGCATATCGAGGCGGTTCATTCGAACTGCACATGGA	1707
721	DB		GCTGTCCAAGGTGCAGCAGGTGCAGCGCGGCATTCGAGGCGGTTCATTCGAACTGCACATGGA	780
1708	QY		CGAAATTTACACTTCGAGGAGCAGCTTTGCAGCAGAGATCCGAGCGTTTCATCTCTCCCAA	1767
781	DB		CGAAATTTATCACTTCGAAGGAGACCTTTGCCAGCAGATCCGAGCGTTTCATCTCTCCCAA	840
1768	QY		GGCAGAGGTGTGCGTGGCGAAACCATGTCCAGCCCTACATCCCATTCCTTGAGGCGCT	1827
841	DB		GGCAGAGGTGTGCGTGGCGAAACCATGTCCAGCCCTACATCCCATTCCTTGAGGCGCT	900
1828	QY		GATGGTCCCAACGACGACGAGGCTTCACTGAGGTGCGAGATGTCTTCTCAAGGAGGTCAAC	1887
901	DB		GATGGTCCCAACGACGAGGCTTCACTGAGGTGCGAGATGTCTTCTCAAGGAGGTCAAC	960
1898	QY		GGACNTGAACCTGAACGTTCATCAACGAGGGCGGATTTGACAAAGCTGGCGAGTACATGGA	1947
961	DB		GGACNTGAACCTGAACGTTCATCAACGAGGGCGGATTTGACAAAGCTGGCGAGTACATGGA	1020
1948	QY		GAACTGTCTCCGGCTGGCGTACACCCCTGAGATCCAGAGCTGCTATGAGAGATGGA	2007
1021	DB		GAACTGTCTCCGGCTGGCGTACACCCCTGAGATCCAGAGCTGCTATGAGAGATGGA	1080
2008	QY		GTCCCTGCAGCTGAGAGGGCTGCAGCAGCGATTTGATGTCTCCAGCAGCTCCGTTCAA	2067
1081	DB		GTCCCTGCAGCTGAGAGGGCTGCAGCAGCGATTTGATGTCTCCAGCAGCTCCGTTCAA	1140
2068	QY		GCACGAGCCCAAGATCCACATGCGGAGGAAATGGAACAATGCCGTGTATAGTTTCGAGAC	2127
1141	DB		GCACGAGCCCAAGATCCACATGCGGAGGAAATGGAACAATGCCGTGTATAGTTTCGAGAC	1200
2128	QY		CCTCTCCACAGGAGCTGGGAGGGGCCCAACGAAGGAGGAGCTGTGCAAGTCCATCCA	2187
1201	DB		CCTCTCCACAGGAGCTGGGAGGGGCCCAACGAAGGAGGAGCTGTGCAAGTCCATCCA	1260
2188	QY		GCGGTCTCGAGCGGGTCTGAAAAAATACCACTACGACGAGAGCTGTGTCGGAAGAG	2247
1261	DB		GCGGTCTCGAGCGGGTCTGAAAAAATACCACTACGACGAGAGCTGTGTCGGAAGAG	1320
2248	QY		GTTCTTCGGGAGGCGCTGTCGAGATCAAGATCCCGTTCTGTCAGAAAGCTGGCCCC	2307
1321	DB		GTTCTTCGGGAGGCGCTGTCGAGATCAAGATCCCGTTCTGTCAGAAAGCTGGCCCC	1380
2308	QY		TACCTGCAAGTTCGAGCTGCCCGGTTCCAGAGAGCTGATCTTCGAGGACTTTGCCAGGTT	2367
1381	DB		TACCTGCAAGTTCGAGCTGCCCGGTTCCAGAGAGCTGATCTTCGAGGACTTTGCCAGGTT	1440
2368	QY		CATCTCGTGGAACCAACGATGAGGAGGTGTGTGTCAGAGCCTGTACGAGGAGCATCCT	2427
1441	DB		CATCTCGTGGAACCAACGATGAGGAGGTGTGTGTCAGAGCCTGTACGAGGAGCATCCT	1500
2428	QY		GAGGCTGTGAAGAGGCGCGGTGTCAGAGGAGAGCAACCTCTACCGGGACAGATGGT	2487
1501	DB		GAGGCTGTGAAGAGGCGCGGTGTCAGAGGAGAGCAACCTCTACCGGGACAGATGGT	1560
2488	QY		CATGCACAAACGAGCCCAACCTGTGCGCGAGGGCGCCCATCGACTGGGG	2547
1561	DB		CATGCACAAACGAGCCCAACCTGTGCGCGAGGGCGCCCATCGACTGGGG	1620
2548	QY		CGAGGAGTACAGCAACGCGGGGGCGGCGCCAGCCAGCCCGGAGTCAAG	2607
1621	DB		CGAGGAGTACAGCAACGCGGGGGCGGCGCCAGCCAGCCCGGAGTCAAG	1680
2608	QY		CACCTCTCGGAAAGCGAGCGCGCGCAAGCAGGTGTCTGTGTCCAGAGTACGGA	2667
1681	DB		CACCTCTCGGAAAGCGAGCGCGCGCAAGCAGGTGTCTGTGTCCAGAGTACGGA	1740
2668	QY		GGTGGGCTGCGCTTTGAGGCTAGCCCTGAGTCAACCACTGCGTCCCGGACGSGT	2727

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 25, 2006, 11:23:00 ; Search time 2583 Seconds
(without alignments)
12235.851 Million cell updates/sec

Title: US-10-612-318-3
Perfect score: 4530
Sequence: 1 gggtccctgcacaaatgcg.....atttaaaaaaaaaaaaaa 4533

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5244320 segs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: N_Geneseq18.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*
- 15: Geneseqn2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4530	100.0	4533	12	ADL23467 Human STR
2	4430	97.8	4623	12	ADL23465 Human STR
3	3763.4	83.1	3768	3	AAC81955 Human Meg
4	3643.4	80.4	3995	5	AAI68201 Human ost
5	3640	80.4	3817	13	ACN38947 Tumour-as
6	3639.8	80.3	4096	5	ADL45287 Human ova
7	3553.4	78.4	3656	13	ADR24468 Breast ca
8	2983.6	65.9	3749	6	ABQ93412 Human CDN
9	1950.8	43.1	2041	4	AAH18056 Human sec
10	1658.6	36.6	1777	3	AAC60052 Human sec
11	1658.6	36.6	1777	8	ADA97962 Human sec
12	1658.6	36.6	1777	8	ADA43868 Human sec
13	1658.6	36.6	1777	10	ADC20118 Human sec
14	1658.6	36.6	1777	10	ADF10578 Human sec
15	1342.6	29.6	1446	2	AAV84420 Human sec
16	1342.6	29.6	1446	4	ABA83203 Human sec
17	1342.6	29.6	1446	9	ACH04704 Novel hum
18	1342.6	29.6	1446	9	ACD44514 Human CDN

19	858	18.9	924	4	AAK56817	Aak56817 Human imm
20	846	18.7	922	3	AAA23460	Aaa23460.cDNA enco
21	648.8	14.3	713	5	AAF68285	Aaf68285 Human lun
22	648.8	14.3	713	6	ABK38196	Abk38196.cDNA enco
23	648.8	14.3	713	8	ACA10525	Aca10525 Human lun
24	648.8	14.3	713	8	ABX99476	Abx99476 Lung can
25	648.8	14.3	713	10	ADH45722	Adh45722 Human lun
26	648.8	14.3	713	12	ADE72259	Ade72259 Human lun
27	648.8	14.3	713	13	ADJ19641	Adj19641 Human lun
28	625.2	13.8	783	4	AAH07715	Aah07715 Human CDN
29	556.4	12.3	578	5	ADL39034	Adl39034 Human ova
30	534	11.8	1074	2	AAK82084	Aax82084 Human SIG
31	457.4	10.1	470	5	ADI73797	Adi73797 Human ova
32	457.4	10.1	470	5	ADI67404	Adi67404 Human ova
33	428.8	9.5	456	5	ADL40423	Adl40423 Human ova
34	408	9.0	434	14	ACL62507	Acl62507 Human col
35	375	8.3	393	5	AAF64572	Aaf64572 Novel hum
36	368.4	8.1	3467	14	ARE04897	Aee04897 Cancer-as
37	351	7.7	2787	15	ABG07464	Aeg07464 Human acc
38	351	7.7	4475	4	AAF54990	Aaf54990 Nucleotid
39	351	7.7	4592	5	ABV23100	Abv23100 Human pro
40	351	7.7	4592	5	ABV28938	Abv28938 Human pro
41	351	7.7	4951	14	AAE04900	Aee04900 Cancer-as
42	351	7.7	6319	14	ADY17310	Ady17310 DNA enco
43	351	7.7	6919	14	ADY17312	Ady17312 DNA enco
44	351	7.7	7243	14	AAE04902	Aee04902 Cancer-as
45	349.4	7.7	3714	4	AAH15758	Aah15758 Human CDN

ALIGNMENTS

RESULT 1
ID ADL23467 standard; CDNA; 4533 BP.
XX ADL23467;
XX AC ADL23467;
XX DT 01-JUL-2004 (first entry)
XX DE Human STR50 short variant CDNA.
XX KW Human, STR50; gene, ss; short variant; neurotoxic stress;
XX KW neurodegenerative disease; stroke; Parkinson's disease;
XX KW Alzheimer's disease; Huntington's disease; tumour; cancer;
XX KW cerebroprotective; vasotropic; antiparkinsonian; neuroprotective;
XX KW nootropic; anticonvulsant.
XX OS Homo sapiens.
XX FT Key Location/Qualifiers
XX FT CDS 644..3109
XX FT FT /*tag= a
XX FT FT /product= "Human STR50 short variant"
XX PN US2004068098-A1.
XX PD 08-APR-2004.
XX PF 01-JUL-2003; 2003US-00612318.
XX PR 01-JUL-2002; 2002US-0393251P.
XX PA (KACH/) KACHALSKY S G.
XX PA (FAER/) FAERMAN A.
XX PA (PELO/) PEL-OR Y.
XX PI Kachalsky SG, Faerman A, Pel-Or Y;
XX DR WPI; 2004-304622/28.
XX DR P-PSDB; ADL23468.
XX FT New STR50 polynucleotides and polypeptides, useful in diagnosing and

Wed Jun 28 10:20:23 2006

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 25, 2006, 11:28:11 ; Search time 24735 Seconds
(without alignments)
11719.170 Million cell updates/sec

Title: US-10-612-318-3
Perfect score: 4530
Sequence: 1 gggctccctgcacaaatgcg.....atttaaaaaaaaaaaaaa 4533

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenBank:

1: gb_env.*

2: gb_pat.*

3: gb_ph.*

4: gb_pl.*

5: gb_pr.*

6: gb_ro.*

7: gb_sts.*

8: gb_sy.*

9: gb_un.*

10: gb_vi.*

11: gb_ov.*

12: gb_htg.*

13: gb_in.*

14: gb_om.*

15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB ID	Description
1	3763.4	83.1	3768	5	AF151783	Homo sapi
2	3643.4	80.4	3995	2	BD016436	Gene deri
3	3643.4	80.4	3995	5	AF192911	Homo sapi
4	3639.8	80.3	4096	2	CQ412106	Sequence
5	3623.2	80.0	3725	2	BD183401	Novel gen
6	3623.2	80.0	3725	5	AB210016	Homo sapi
7	3623.2	80.0	3725	5	AB210016	Human nuc
8	3623.2	80.0	3725	5	AB210016	Human nuc
9	3623.2	80.0	3725	5	AB210016	Human nuc
10	3623.2	80.0	3725	5	AB210016	Human nuc
11	3623.2	80.0	3725	5	AB210016	Human nuc
12	3623.2	80.0	3725	5	AB210016	Human nuc
13	3623.2	80.0	3725	5	AB210016	Human nuc
14	3623.2	80.0	3725	5	AB210016	Human nuc
15	3623.2	80.0	3725	5	AB210016	Human nuc
16	3623.2	80.0	3725	5	AB210016	Human nuc
17	3623.2	80.0	3725	5	AB210016	Human nuc
18	3623.2	80.0	3725	5	AB210016	Human nuc

AK035951 Homo sapi
BD192925 207 human
CQ821921 Sequence
AR693816 Sequence
BC024639 Mus muscu
BC104682 Rattus no
CT030569 Xenopus t
BD270076 Secreted
AL390116 Human DNA
BX511185 Homo sapi
BC072979 Xenopus l
CQ716711 Sequence
AR272491 Sequence
AR276072 Sequence
AR406347 Sequence
AR440197 Sequence
AR472355 Sequence
AR543008 Sequence
AR638166 Sequence
AR694839 Sequence
AR707426 Sequence
AX062576 Sequence
AX367493 Sequence
BD149707 Primer fo
AX869645 Sequence
AC135968 Homo sapi
CQ405853 Sequence

ALIGNMENTS

RESULT 1
AF151783 3768 bp mRNA linear PRI 30-MAY-2001
LOCUS Homo sapiens MEG3 (MEG3) mRNA, complete cds.
DEFINITION AF151783
ACCESSION AF151783
VERSION AF151783.1 GI:14248494
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (Bases 1 to 3768)
AUTHORS Miyata, T., Inagi, R., Yasuda, Y. and Kurokawa, K.
TITLE Homo sapiens meg-3 mRNA, complete cds
JOURNAL Unpublished
REFERENCE 2 (Bases 1 to 3768)
AUTHORS Miyata, T., Inagi, R., Yasuda, Y. and Kurokawa, K.
TITLE Direct Submission
JOURNAL Submitted (17-MAY-1999) Institute of Medical Sciences, Department of Internal Medicine, Tokai University School of Medicine, Bohseidai, Isehara 259-1193, Japan

FEATURES

source

1. 3768
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

gene

1. 3768
/gene="MEG3"

CDS

53..2254
/gene="MEG3"
/codon_start=1
/product="MEG3"
/protein_id="AAK57556.1"
/db_xref="GI:14248495"

/translation="MGWGEKTKILPELFQYEDQYGVLFNSMRHEIEGTGLPQAO
LLRWKPLDRIIVFSGNLFQHQDSKWRNRSFLPHNYGLVLYENKAAVROVPPRAQ
VINSAGYKLTSDVQYLEIGNSLPGTTAKSGAPILKCTQFPFLILHWPVARYVFC
NMTEAQDKQWALQDCIRHCNNGIPEDSKVGFAPFTDAIRNRSKLELGTWEMLCG
NEVCILLNLVMEELGPELKGPKQERQWQIQISDAVYHVMYEQAKRFE
VLKSVQVQVQPMQAVIRTDMDQIITIKELLAKIRAFILPKAEVCRNRHVQPIPSIL

EALMVTSPGQFTEVRDVFKEVTTMNLVINEGGIDKLGEVNEKLSRLAYHPLKMQSC
YKMSRLDGLQORFDVSSVSFKQRAQTHMRQEDNAVYTFETLLHQLGKGPKE
ELCKSTQRLVRLVKYDYSVSFKRFFREALQISIFLKLAPTKCKSLPRFOE
LIFEDARFIVENTYEVVLOTVMKDIILQAVKEAAVQRKENVYRDSVMHNSDPLNH
LLAEGAPIDWEXNSGGGSPSPSPSATSLEKRRRAKQVVSVVQDEVLGPFEA
SPSPSPSPDGVTEIRGLAQLRPESPPAPILLGAPAGESPOPKAAPAS9PPA
SPLQHLHPKAVDLOPPKPSDQETGEQVSPSSHPALHTTTEDSAGVQTEF"

QY	1666	GGTGACGCGCCATGACGAGCGCTCATCCGAACTGACATGGAACCAAAATATACCTCCAA	1725
DB	901	GGTGACGCGCCATGACGAGCGCGTCTCCGAACTGACATGGAACCAAAATATACCTCCAA	960
QY	1726	GGAGCACTTGGCAGCAAGATGCGAGCGCTTCTCTCCCAAGGAGGAGGTGCGTGGG	1785
DB	961	GGAGCTCTTGGCAGCAAGATGCGAGCGCTTCTCTCCCAAGGAGGAGGTGCGTGGG	1020
QY	1786	GAACATGTTCAGCGCTACATCCATCTCGAGGCGCTGTGTTGCCACACGCA	1845
DB	1021	GAACATGTTCAGCGCTACATCCATCTCGAGGCGCTGTGTTGCCACACGCA	1080
QY	1846	GGGCTTCACTGAGGTGCGAGATGTCTTCTTAAGGAGGTGCGAGCATGACCTGACCT	1905
DB	1081	GGGCTTCACTGAGGTGCGAGATGTCTTCTTAAGGAGGTGCGAGCATGACCTGACCT	1140
QY	1906	CATCAAGGAGGCGCATTTGAAAGCTGGCGGAGTACATGGAAGAGTGTCTCCGGCTGGC	1965
DB	1141	CATCAAGGAGGCGCATTTGAAAGCTGGCGGAGTACATGGAAGAGTGTCTCCGGCTGGC	1200
QY	1966	GTACACCCCTGAAGATGCGAGCTGTATGAAGAGATGAGAGTGTCTGCGACTGGACGG	2025
DB	1201	GTACACCCCTGAAGATGCGAGCTGTATGAAGAGATGAGAGTGTCTGCGACTGGACGG	1260
QY	2026	GCTCAGCAGCGATTTGATGTGTCCAGCAGTCCGTGTTCAAGCAGCAGGCGCAGATCCA	2085
DB	1261	GCTCAGCAGCGATTTGATGTGTCCAGCAGTCCGTGTTCAAGCAGCAGGCGCAGATCCA	1320
QY	2086	CATCGGGGAGCAAAATGGAATGCGGTGTATGCTTGGAGAGCTCTCTGCGACGAGAGCT	2145
DB	1321	CATCGGGGAGCAAAATGGAATGCGGTGTATGCTTGGAGAGCTCTCTGCGACGAGAGCT	1380
QY	2146	GGGGAAGGCGCCACCAAGGAGGAGTGTGCAAGTCCATCCAGCGGCTCTTGGAGCGGGT	2205
DB	1381	GGGGAAGGCGCCACCAAGGAGGAGTGTGCAAGTCCATCCAGCGGCTCTTGGAGCGGGT	1440
QY	2206	GCTGAAAAAATACGACTACGACAGAGCTGTGTCGCGAAGAGGTTCTTCCGGAGGCGCT	2265
DB	1441	GCTGAAAAAATACGACTACGACAGAGCTGTGTCGCGAAGAGGTTCTTCCGGAGGCGCT	1500
QY	2266	GCTGAGATCAGCATCCCGTCTCTGCTCAAGAGAGTGGCCCTTACTGCAAGTGGAGCT	2325
DB	1501	GCTGAGATCAGCATCCCGTCTCTGCTCAAGAGAGTGGCCCTTACTGCAAGTGGAGCT	1560
QY	2326	GCCCCGTTCACGAGCTGATCTTCGAGGACTTTGCGAGGTTATCTCTGTTGGAAAAAC	2385
DB	1561	GCCCCGTTCACGAGCTGATCTTCGAGGACTTTGCGAGGTTATCTCTGTTGGAAAAAC	1620
QY	2386	GTAAGAGAGGTGCTGCGACCGTCTAGAGGAGTCTCTGAGGCTGTGAAGGAGGC	2445
DB	1621	GTAAGAGAGGTGCTGCGACCGTCTAGAGGAGTCTCTGAGGCTGTGAAGGAGGC	1680
QY	2446	CGCGGTGAGAGGAGGAGCAACCTCTACCGGAGCAGCATGTTGTCATGCAACAGCGACC	2505
DB	1681	CGCGGTGAGAGGAGGAGCAACCTCTACCGGAGCAGCATGTTGTCATGCAACAGCGACC	1740
QY	2506	CAACCTGCACTGTGCGAGGCGCCCCATCGACTGGGCGAGGAGTACAGCAACAG	2565
DB	1741	CAACCTGCACTGTGCGAGGCGCCCCATCGACTGGGCGAGGAGTACAGCAACAG	1800
QY	2566	CGGCGGGGCGGAGCGCCAGCAACCCCGAGTACAGCACCCTCTCGGAAAAAGCG	2625
DB	1801	CGGCGGGGCGGAGCGCCAGCAACCCCGAGTACAGCACCCTCTCGGAAAAAGCG	1860
QY	2626	ACGGCGCGCCAGAGGTGTCTGTGTGTCAGGATGAGAGGTGGGCTGCTCTTTGA	2685
DB	1861	ACGGCGCGCCAGAGGTGTCTGTGTGTCAGGATGAGAGGTGGGCTGCTCTTTGA	1920
QY	2686	GGCTAGCGCTGAGTACCAACCTGCGTCCCGGAGCGGTGTCACTGAGATCCGAGGCT	2745
DB	1921	GGCTAGCGCTGAGTACCAACCTGCGTCCCGGAGCGGTGTCACTGAGATCCGAGGCT	1980
QY	2746	GCTGCGCCCAAGGTCTGCGGCTGAGAGGCGCCCAACAGCGCGCCCTCTCTCAAGCGGCG	2805

ORIGIN

Query Match	83.1%;	Score	3763.4;	DB 5;	Length	3768;			
Best Local Similarity	99.9%;	Pred. No.	0;						
Matches	3767;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	766	CAGGAACTGGGGCAGCTCCGGTCCCTTCTCTTTTGGGGCTCTCACTCTGGAGATGGGGTG	825						
DB	1	CAGGAACCTGGGCCAGCTCCGGTCCCTTCTCTTTTGGGGCTCTCACTCTGGAGATGGGGTG	60						
QY	826	GATGGGAGAAAAAACCGGAGATCTGACGGAGTTCCTCCAGTTCATGAAGACCACTA	885						
DB	61	GATGGGAGAAAAAACCGGAGATCTGACGGAGTTCCTCCAGTTCATGAAGACCACTA	120						
QY	886	TGGCGTGGCTCTCTTCAACAGCATGCGCCATGAGATTGAGGCGACGGGGCTGCCAGGC	945						
DB	121	TGGCGTGGCTCTCTTCAACAGCATGCGCCATGAGATTGAGGCGACGGGGCTGCCAGGC	180						
QY	946	CCAGCTGCTCTGGCGAAGGTGCGCATGAGCAGCGCATCGTCTTCTCGGGGAACCTCTT	1005						
DB	181	CCAGCTGCTCTGGCGAAGGTGCGCATGAGCAGCGCATCGTCTTCTCGGGGAACCTCTT	240						
QY	1006	CCAGCACCCAGGAGGACAGCAAGATGGAGAAACCGCTTTCAGCTGTGCCCCCAACTA	1065						
DB	241	CCAGCACCCAGGAGGACAGCAAGATGGAGAAACCGCTTTCAGCTGTGCCCCCAACTA	300						
QY	1066	CGGGCTGGTCTTACGAAAAAAGCGGCTTATGAGCGGCGAGTCCACACAGAGCCGT	1125						
DB	301	CGGGCTGGTCTTACGAAAAAAGCGGCTTATGAGCGGCGAGTCCACACAGAGCCGT	360						
QY	1126	CATCAACAGTGCAGGCTACAAATCTCTCAGCTCCGTGGACCAATACCTGGAGCTCATTTGG	1185						
DB	361	CATCAACAGTGCAGGCTACAAATCTCTCAGCTCCGTGGACCAATACCTGGAGCTCATTTGG	420						
QY	1186	CAACTCTTACAGGACACGCAAGTCCGGCAGTGGCCCCCATCTCAAGTGGCCCCAC	1245						
DB	421	CAACTCTTACAGGACACGCAAGTCCGGCAGTGGCCCCCATCTCAAGTGGCCCCAC	480						
QY	1246	ACAGTTCGCCGCTCATCTCTGCGCATCTTATGCGCGTCACTACTACTTCTGCAATGATAC	1305						
DB	481	ACAGTTCGCCGCTCATCTCTGCGCATCTTATGCGCGTCACTACTACTTCTGCAATGATAC	540						
QY	1306	AGAAGCCGACAGGACCAAGTGGCAGGCTGTGTCGAGGAGTGTGATCCGGCACTGCAACAA	1365						
DB	541	AGAAGCCGACAGGACCAAGTGGCAGGCTGTGTCGAGGAGTGTGATCCGGCACTGCAACAA	600						
QY	1366	TGGAATCCCTGAGGACTTCCAAGGTAGAGGGCCCTCGGTTTCAAGATGCCATTCGCAATGA	1425						
DB	601	TGGAATCCCTGAGGACTTCCAAGGTAGAGGGCCCTCGGTTTCAAGATGCCATTCGCAATGA	660						
QY	1426	CCGACAGTCCAGGAGCTGTACGGCACTGGGAGATGCTGTGTGGAAACGAGGTGCAGAT	1485						
DB	661	CCGACAGTCCAGGAGCTGTACGGCACTGGGAGATGCTGTGTGGAAACGAGGTGCAGAT	720						
QY	1486	CCTGAGCAACTGTGTGATGAGGAGCTGGGCGCTTGTGAGTGAAGGAGAGCTTCGGCCCGCG	1545						
DB	721	CCTGAGCAACTGTGTGATGAGGAGCTGGGCGCTTGTGAGTGAAGGAGAGCTTCGGCCCGCG	780						
QY	1546	GCTGAAGGGGAAAACCGCAGGAGCGGACCGGAGTGTGATCCAGTCTCGGAGCGCTGTA	1605						
DB	781	GCTGAAGGGGAAAACCGCAGGAGCGGACCGGAGTGTGATCCAGTCTCGGAGCGCTGTA	840						
QY	1606	CCACATGTGTACGAGCAGGCGCAAGCGCGCTTCCAGGAGGTGTGTTTCAAGGTGCAGCA	1665						
DB	841	CCACATGTGTACGAGCAGGCGCAAGCGCGCTTCCAGGAGGTGTGTTTCAAGGTGCAGCA	900						

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2006, 01:56:20 ; Search time 37.2928 Seconds
(without alignments)
1856.571 Million cell updates/sec

Title: US-10-612-318-4
Perfect score: 4139
Sequence: 1 MEEVGRWGLKPRLAGLRG.....SSHPALHTTSDSAGVQTEF 791

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : Issued Patents AA.*
- 1: /EMC/Celerra_SIDS3/ptodata/2/iaa/5 COMB.pcp.*
 - 2: /EMC/Celerra_SIDS3/ptodata/2/iaa/6 COMB.pcp.*
 - 3: /EMC/Celerra_SIDS3/ptodata/2/iaa/7 COMB.pcp.*
 - 4: /EMC/Celerra_SIDS3/ptodata/2/iaa/H COMB.pcp.*
 - 5: /EMC/Celerra_SIDS3/ptodata/2/iaa/PCTUS COMB.pcp.*
 - 6: /EMC/Celerra_SIDS3/ptodata/2/iaa/RE COMB.pcp.*
 - 7: /EMC/Celerra_SIDS3/ptodata/2/iaa/backfiles.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	176.5	4.3	2090	2	US-09-538-092-1081
2	176.5	4.3	2120	2	US-09-949-016-9768
3	159.5	3.9	1248	1	US-09-080-897-2
4	159.5	3.9	1248	2	US-09-323-735-2
5	158.5	3.8	1315	2	US-08-899-595-3
6	152	3.7	864	2	US-09-538-092-1268
7	151	3.6	1255	1	US-09-080-897-4
8	151	3.6	1255	2	US-08-899-595-1
9	151	3.6	1255	2	US-09-323-735-4
10	147	3.6	688	2	US-09-141-047-8
11	147	3.6	1274	2	US-09-095-443-2
12	147	3.6	1636	3	US-10-433-794-1
13	146.5	3.5	693	2	US-09-081-385-154
14	146.5	3.5	693	2	US-09-752-639-154
15	146.5	3.5	693	2	US-09-712-813-154
16	146.5	3.5	693	2	US-09-700-354A-154
17	144	3.5	983	2	US-09-311-236-2
18	144	3.5	983	2	US-09-912-917-2
19	144	3.5	1444	2	US-09-902-540-16727
20	143.5	3.5	916	2	US-09-949-016-11417
21	142.5	3.4	915	2	US-09-538-092-863
22	142.5	3.4	916	2	US-09-949-016-6611
23	142.5	3.4	1020	2	US-09-538-092-911
24	139.5	3.4	638	2	US-09-949-016-10708
25	137.5	3.3	912	5	PCT-US95-03747-2
26	137.5	3.3	1231	2	US-08-714-741-41

27	136.5	3.3	2482	1	US-08-328-254-6
28	136.5	3.3	3210	2	US-09-538-092-1154
29	136	3.3	603	2	US-09-906-779-4
30	136	3.3	694	2	US-09-949-016-7117
31	134	3.2	641	2	US-08-961-083-160
32	134	3.2	641	2	US-09-536-784-160
33	134	3.2	641	2	US-09-765-271-160
34	134	3.2	641	2	US-09-765-272A-160
35	134	3.2	863	2	US-10-067-457-5
36	133	3.2	868	2	US-09-800-729-106
37	133	3.2	921	2	US-09-800-729-199
38	132.5	3.2	1388	1	US-08-685-576-1
39	132	3.2	802	2	US-09-823-240A-2
40	132	3.2	1312	2	US-09-041-886-19
41	132	3.2	1312	2	US-09-648-281-2
42	132	3.2	1312	2	US-09-707-919A-19
43	132	3.2	1312	2	US-09-083-268-3
44	132	3.2	1312	2	US-08-981-998A-3
45	131.5	3.2	885	2	US-09-914-259-10

ALIGNMENTS

RESULT 1
US-09-538-092-1081
; Sequence 1081, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqformatter Version 0.9
; SEQ ID NO 1081
; LENGTH: 2090
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)-(0)
; OTHER INFORMATION: Polypeptide Accession Number P35658
US-09-538-092-1081

Query Match	4.3%	Score 176.5;	DB 2;	Length 2090;
Best Local Similarity	19.6%;	Pred. No. 3.5e-05;		
Matches	148;	Conservative	126;	Mismatches 258; Indels 225; Gaps 38;
Qy	154	ROVPPRAVINSAGYKI-----LTSVDQYLELIGNSLP---GTTAKSGSPILKCP	200	
Db	604	QSPAPMPPSSASGKPAASGPLSGHTPLSAPPSSVPLKSSVLPSPSGRGAQSSSPVPSMV	663	
Qy	201	TOFPLIL-----WHPYARHYFCWMTAEQDKWQ-----AVLODCIRHCNNGIPE-DS	247	
Db	664	OKSPRIITPPAAKPGSPQAKSLQ-PAVAEKQGHOMKSDPVMAGIGETIAHFKQELKKA	722	
Qy	248	KVEGPAF-----TAIRMYRQSK-----ELYGTWMLCGNEVQITLSNLVMEELGPELK	295	
Db	723	RISKACFOGVTSEMWMLETSDDDLHTLLEIKETTESLHGDITSLKLTLLGFAGVEEA	782	
Qy	296	AELGPRKKGKPOQRQOWIQISDAMY-HMVYE-----QAKARFEVLSKVQOVOPAMQAV	349	
Db	783	RE-----QNERNR-----DSGYLHLLYKRLDPKSEAQLQEIIRLHQYVKFAVQ--	826	
Qy	350	INTMDQIITSK--EHLASKIRAFILPKAEVCVNRHVQYIPISLEALMWPTSQGFTFVR	407	

753 RTSKACFQVGTSEEMKMLRTESDDLTLLEIKETTESLHGDISLTKITLLEGFAGVEEA 812
296 AELGPRLKGPORQOWIQISDAVY-HMVE-----QAKARFEVLSKVQVQPMQAV 349
813 RE-----QNERNR-----DSGYLHLYKRPDLPKSEAQLQIRLHLYVQKFAVQ-- 856
350 IRTDMDQITSK--EHLASKIRAFILPKAECVVRNHOVQYIPSIIEALMVPTSGFTVEVR 407
857 ---DVNDVLDLEWDQHLQKKK-----QRR-----LLVPE-----R 884
408 DVFFKEVTDNMLNVINEGGIDKLGEYMEKLSRLAYHPLKMQSCYKEMESLRLDGLQORFD 467
885 ETLFNTLAN-NREIN-----QQRKRLNHL-----VDSLQQLRL 917
468 VSSVTSVFKQRAQHMRQMDNAVYTFETLLHQLGKPTKEELCKSIORVLERVLKYYD 527
918 YKQISLMSLSAV-----PSQSIHSFSDSL-----ESLCNAL---LKTIESHTK 960
528 DSSSVKRF--FREALLQISIPFLKLAFTCKSELP-----RFQELIFEDFAR 574
961 SLPKVPKAKSPMQAQLR---NFLAKRKTTPVRSTAPASLSRSASFLSORYYEDL-DEVSS 1016
575 FILVENTYEEVVLQTVMKD---LLOAVKEAAVORKHNL-----YRDSMWMHNSDP 621
1017 TSSVSQSLSESDARTCKDDEAVVQAPRHAPVVRTPSIQPSLLPHAAPAKSHLVHGSSP 1076
622 NL-----HLLAEGAPIDWGEYNSGGSGSPSPSTESA----- 655
1077 GWGTSVATSASKIIPQGA--DSTMLATKTVKHGAPSPSHPIAPQAAAAARRQMASQ 1134
656 ---TLSEKRRRAKQVSVVQDEEV-----GLPEASPESSPPSPASPDGVT 697
1135 APAVNTLTES--TLKNVPQVNVQELKNNPATPTAMGSSVPYSTAKTTPHVLTPVAANQ 1192
698 IR-GLLQGLRPESPPA-GPLNGAPAGESPQKAA-----PEASSPPASPLQHLPPKA 751
1193 AKQGLSLNLSKPGFTPASQOLSSGDKASGTAKIETAVTSTPSASQGFSPSPSGTG 1252
752 VDLG--PPKPSDQETGEQVSSPSHPALHTTTEDSAG 786
1253 FNFGIITPSSNFTAAQGATPTKESQPDADFSSGG 1289

RESULT 3
US-09-080-897-2
; Sequence 2, Application US/09080897
; Patent No. 5985574
; GENERAL INFORMATION:
; APPLICANT: King, Mary-Claire
; APPLICANT: Lynch, Eric D.
; APPLICANT: Lee, Ming
; APPLICANT: Morrow, Jan E.
; APPLICANT: Welcsh, Piri L.
; APPLICANT: Leon, Pedro E.
; TITLE OF INVENTION: Modulators of Actin
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/080,897
; FILING DATE:
; CLASSIFICATION:

827 ---DVNDVLDLEWDQHLQKKK-----QRR-----LLVPE-----R 854
408 DVFFKEVTDNMLNVINEGGIDKLGEYMEKLSRLAYHPLKMQSCYKEMESLRLDGLQORFD 467
855 ETLFNTLAN-NREIN-----QQRKRLNHL-----VDSLQQLRL 887
468 VSSVTSVFKQRAQHMRQMDNAVYTFETLLHQLGKPTKEELCKSIORVLERVLKYYD 527
888 YKQISLMSLSAV-----PSQSIHSFSDSL-----ESLCNAL---LKTIESHTK 930
528 DSSSVKRF--FREALLQISIPFLKLAFTCKSELP-----RFQELIFEDFAR 574
931 SLPKVPKAKSPMQAQLR---NFLAKRKTTPVRSTAPASLSRSASFLSORYYEDL-DEVSS 986
575 FILVENTYEEVVLQTVMKD---LLOAVKEAAVORKHNL-----YRDSMWMHNSDP 621
987 TSSVSQSLSESDARTCKDDEAVVQAPRHAPVVRTPSIQPSLLPHAAPAKSHLVHGSSP 1046
622 NL-----HLLAEGAPIDWGEYNSGGSGSPSPSTESA----- 655
1047 GWGTSVATSASKIIPQGA--DSTMLATKTVKHGAPSPSHPIAPQAAAAARRQMASQ 1104
656 ---TLSEKRRRAKQVSVVQDEEV-----GLPEASPESSPPSPASPDGVT 697
1105 APAVNTLTES--TLKNVPQVNVQELKNNPATPTAMGSSVPYSTAKTTPHVLTPVAANQ 1162
698 IR-GLLQGLRPESPPA-GPLNGAPAGESPQKAA-----PEASSPPASPLQHLPPKA 751
1163 AKQGLSLNLSKPGFTPASQOLSSGDKASGTAKIETAVTSTPSASQGFSPSPSGTG 1222
752 VDLG--PPKPSDQETGEQVSSPSHPALHTTTEDSAG 786
1223 FNFGIITPSSNFTAAQGATPTKESQPDADFSSGG 1259

RESULT 2
US-09-949-016-9768
; Sequence 9768, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9768
; LENGTH: 2120
; TYPE: PRI
; ORGANISM: Human
US-09-949-016-9768

Query Match 4.3%; Score 176.5; DB 2; Length 2120;
Best Local Similarity 19.6%; Pred. No. 3.6e-05;
Matches 148; Conservative 126; Mismatches 258; Indels 225; Gaps 38;

QY 154 RQVPRVINSAGYKI-----LTSVDQVLELIGNSLP---GTTAKSGSAPILKCP 200
DB 634 QSAPMPSPFSSAKPAASGFLSHPTLSAPPSSVPLKSSVLSPSGRAGQSSSPVSMV 693
QY 201 TOFPLIL-----WHPYARHYFCMWTAEQDKWQ-----AVLQDCIRHCNNGIPE-DS 247
DB 694 QKSPRIPTPAKPGSPQAKSLQ-PAVAEKQHQWKSDFPMAGIGEEIAHAFQKELELKA 752
QY 248 KVEGPAF----TDAIRMYRSK-----ELYGTWMLCGNEVQILSNLWYEEGLPELK 295

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2006, 01:45:00 ; Search time 183.029 seconds
(without alignments)
3997.655 Million cell updates/sec

Title: US-10-612-318-4
Perfect score: 4139
Sequence: 1 MEEVGRWGLKRLAGLRG.....SSHPALHTTSDSAGVQTEF 791

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Lisying first 45 summaries

Database : UniProt 7.2.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3824	92.4	733	1 NIBL HUMAN	Q96tal homo sapien
2	3824	92.4	733	2 Q5VVM6 HUMAN	Q5VVM6 homo sapien
3	3792	91.6	739	2 Q4LE55 HUMAN	Q4LE55 homo sapien
4	3792	91.6	746	2 Q5VVM7 HUMAN	Q5VVM7 homo sapien
5	3433.5	83.0	749	1 NIBL MOUSE	Q8rlf1 mus musculus
6	3433.5	83.0	749	2 Q543S7 MOUSE	Q543S7 m 4 days ne
7	3427.5	82.8	749	2 Q8BQ71 MOUSE	Q8BQ71 mus musculus
8	3426.5	82.8	749	2 Q3U813 MOUSE	Q3U813 mus musculus
9	3424.5	82.7	749	2 Q3UC84 MOUSE	Q3UC84 mus musculus
10	3421.5	82.7	749	2 Q3TCV6 MOUSE	Q3TCV6 mus musculus
11	3421.5	82.7	749	2 Q3UIE4 MOUSE	Q3UIE4 mus musculus
12	3417.5	82.6	749	2 Q3UDW4 MOUSE	Q3UDW4 mus musculus
13	3310.5	80.0	706	2 Q8CC78 MOUSE	Q8CC78 mus musculus
14	2590	62.6	526	2 Q6NMY7 HUMAN	Q6NMY7 homo sapien
15	2187.5	52.9	888	2 Q6GPX6 XENLA	Q6GPX6 xenopus lae
16	2027	49.0	396	2 Q9H8K1 HUMAN	Q9H8K1 homo sapien
17	1943	46.9	631	2 Q4T3H2 TETNG	Q4T3H2 tetraodon n
18	1761.5	42.6	392	2 Q3MHT9 RAT	Q3MHT9 rattus norv
19	1617.5	39.1	700	2 Q4REAL TETNG	Q4REAL tetraodon n
20	1425	34.4	327	2 Q2YD88 HUMAN	Q2YD88 homo sapien
21	1162	28.1	928	2 Q2TTR2 HUMAN	Q2TTR2 homo sapien
22	1160	28.0	928	1 NIBA HUMAN	Q9bzq8 homo sapien
23	1157	28.0	926	2 Q3UW53 MOUSE	Q3UW53 mus musculus
24	1085.5	26.2	845	2 Q6PE79 MOUSE	Q6PE79 mus musculus
25	1048	25.3	201	2 Q9H6L6 HUMAN	Q9H6L6 homo sapien
26	951.5	23.0	795	2 Q5F374 CHICK	Q5F374 gallus gall
27	916.5	22.1	584	2 Q3TDH5 MOUSE	Q3TDH5 mus musculus
28	849	20.5	277	2 Q4T8L1 TETNG	Q4T8L1 tetraodon n
29	827.5	20.0	690	2 Q5RG84 BRARE	Q5RG84 brachydanio
30	818.5	19.8	725	2 Q9ESL7 MOUSE	Q9ESL7 mus musculus
31			736	2 Q9ESN0 RAT	Q9ESN0 rattus norv

32 753.5 18.2 592 2 Q4S226 TETNG Q4S226 tetraodon n
33 530.5 12.8 666 2 Q86XR4 HUMAN Q86XR4 homo sapien
34 522.5 12.6 620 2 Q8N894 HUMAN Q8N894 homo sapien
35 519.5 12.6 697 2 Q86XR2 HUMAN Q86XR2 homo sapien
36 518.5 12.5 651 2 Q86XR3 HUMAN Q86XR3 homo sapien
37 468 11.3 579 2 Q8TEQ3 HUMAN Q8TEQ3 homo sapien
38 432.5 10.4 591 2 Q2NL48 RAT Q2NL48 rattus norv
39 421.5 10.2 596 2 Q2TG66 MOUSE Q2TG66 homo sapien
40 415.5 10.0 510 2 Q3TD68 MOUSE Q3TD68 mus musculus
41 312.5 7.6 400 2 Q3TOH6 MOUSE Q3TOH6 mus musculus
42 256 6.2 528 2 Q55G75 DICDI Q55G75 dicystoteli
43 220 5.3 99 2 Q5TEY9 HUMAN Q5TEY9 homo sapien
44 181 4.4 665 2 Q4Q387 LEIWA Q4Q387 leishmania
45 178.5 4.3 2080 2 Q3KQZ0 HUMAN Q3KQZ0 homo sapien

ALIGNMENTS

RESULT 1
NIBL HUMAN
ID NIBL HUMAN STANDARD; PRT; 733 AA.
AC Q96TAL; Q8BUS2; Q9NT35;
DT 13-AUG-2002, integrated into UniProtKB/Swiss-Prot.
DT 13-AUG-2002, sequence version 2.
DT 07-FEB-2006, entry version 28.
DE Niban-like protein (Meg-3).
GN Name=C9orf88;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
[1]
RN NUCLEOTIDE SEQUENCE [MRNA].
RP Miyata T., Inagi R., Yasuda Y., Kurokawa K.;
RA "Homo sapiens meg-3 mRNA, complete cds."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 38-733.
RC TISSUE=Testis;
RG The German cDNA consortium;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
[3]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 278-733.
RC TISSUE=Cervix;
MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ustin T.B., Toshlyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Bosak S.A., McWen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Pahey J., Helton J.E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Schmutz J., Myers R.M.,
Rodriguez A.C., Grimwood J., Schmutz J., Smailus D.E.,
Butterfield J.S., Krzywinski M.I., Skalska U., Marra M.A.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[4]
RN PHOSPHORYLATION SITES SER-652; SER-679 AND SER-683.
PubMed=15302935; DOI=10.1073/pnas.0404720101;
PubMed=15302935; DOI=10.1073/pnas.0404720101;
Beausoleil S.A., Jedrychowski M., Schwartz D., Elias J.E., Villen J.,
Li J., Cohn M.A., Cantley L.C., Gygi S.P.;
"Large-scale characterization of HeLa cell nuclear phosphoproteins."

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2006, 01:50:05 ; Search time 29.9324 seconds
(without alignments)
2542.646 Million cell updates/sec

Title: US-10-612-318-4
Perfect score: 4139
Sequence: 1 MEEVGRWGLKPLRLAGLRG.....SSHPALHTTSDSAGVQTF 791

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Leaving first 45 summaries

Database :
1: p1r1.*
2: p1r2.*
3: p1r3.*
4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3624	87.6	696	T46394	hypothetical prote
2	176.5	3.9	2090	S26058	probable transform
3	162	3.9	858	S15762	neurofilament trip
4	161.5	3.9	1259	T16038	hypothetical prote
5	156	3.8	1461	T41643	probable involveme
6	152.5	3.7	1087	QPMWSH	neurofilament trip
7	151	3.6	864	A40671	dynam, internal
8	151	3.6	1255	T31065	diaphanous protein
9	148	3.6	1017	PC4035	cell-cyclic-depende
10	147	3.6	990	T14756	hypothetical prote
11	147	3.6	7962	I38346	elastic titin - hu
12	146.5	3.5	1400	T31555	hypothetical prote
13	145.5	3.5	802	S44382	pML protein, splic
14	145.5	3.5	1130	T34081	hypothetical prote
15	145	3.5	1018	T43168	hypothetical prote
16	145	3.5	1072	A37221	neurofilament trip
17	145	3.5	1260	TVRTNU	protein-tyrosine k
18	144.5	3.5	871	T48502	hypothetical prote
19	144	3.5	798	I50479	neurofilament medi
20	143.5	3.5	1251	A56677	neuronal cell cycl
21	142.5	3.4	802	S42518	pML protein, splic
22	142.5	3.4	851	B40671	dynam, internal
23	142.5	3.4	916	A27864	neurofilament trip
24	142.5	3.4	1020	QFHUH	neurofilament trip
25	141	3.4	581	T22341	hypothetical prote
26	139	3.4	2259	S29236	calcium channel pr
27	138.5	3.3	268	H84684	En/Spm-like transp
28	138.5	3.3	1282	JE0120	glycoprotein A - m
29	138.5	3.3	1388	S74245	serine/threonine-s

ALIGNMENTS

RESULT 1

T46394
hypothetical protein DKFp434H0820.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
R:Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23031
A:Accession: T46394
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-696 <AAA>
A:Cross-references: UNIPROT:Q967A1; UNIPARC:UPI000016ACC1; EMBL:AL137555
A:Experimental source: adult testis; clone DKFp434H0820
C:Genetics:
A:Note: DKFp434H0820.1

Query Match 87.6%; Score 3624; DB 2; Length 696;

Best Local Similarity 100.0%; Pred. No. 2.4e-200; Indels 0; Gaps 0;
Matches 696; Conservative 0; Mismatches 0;

Qy	96	TGLPQAQLLRKVP	LDRI	VF	SGNL	FQHQEDSKKWRN	RFS	LVPHNYGLVLYENKAA	YERQ	155
Db	1	TGLPQAQLLRKVP	LDRI	VF	SGNL	FQHQEDSKKWRN	RFS	LVPHNYGLVLYENKAA	YERQ	60
Qy	156	VPPRAVINSAGYK	ILTS	VDQY	LELI	GNSLP	GTAKSGSAPIL	KCPTQFPFLI	WHPYARHY	215
Db	61	VPPRAVINSAGYK	ILTS	VDQY	LELI	GNSLP	GTAKSGSAPIL	KCPTQFPFLI	WHPYARHY	120
Qy	216	YFCWMTAEQDK	QWAV	LODCI	RHCNNGI	PEDSK	VEGPAFTDA	IRMYRQSKELY	GTWEMILC	275
Db	121	YFCWMTAEQDK	QWAV	LODCI	RHCNNGI	PEDSK	VEGPAFTDA	IRMYRQSKELY	GTWEMILC	180
Qy	276	GNEVQILSNLW	MEELGPEL	KAE	LPRLK	GPQR	QWQI	ISDAVYHVMVYEQAKAR	FEV	335
Db	181	GNEVQILSNLW	MEELGPEL	KAE	LPRLK	GPQR	QWQI	ISDAVYHVMVYEQAKAR	FEV	240
Qy	336	LSKVQVQVQPA	QAVIR	TMDQI	ITTS	KEHLASKI	RAFILPKAE	VCVRNHVQPI	PSILEAL	395
Db	241	LSKVQVQVQPA	QAVIR	TMDQI	ITTS	KEHLASKI	RAFILPKAE	VCVRNHVQPI	PSILEAL	300
Qy	396	MVPTSQGTETVR	DV	FKV	VTDMNL	INVEGGIDKLGEYMEK	LSRLAYHPLKMQSCYE	KME		455
Db	301	MVPTSQGTETVR	DV	FKV	VTDMNL	INVEGGIDKLGEYMEK	LSRLAYHPLKMQSCYE	KME		360
Qy	456	SLRLDGLQGR	FDV	SSVTS	VFKORAOIHM	REOMDN	NAVYTFETL	HLHQLGKGP	KEELCKSTQ	515
Db	361	SLRLDGLQGR	FDV	SSVTS	VFKORAOIHM	REOMDN	NAVYTFETL	HLHQLGKGP	KEELCKSTQ	420
Qy	516	RVLSERVLLKYD	VS	SSVSR	KRRFF	REALLQISIP	FLKLLK	LAPTCKSEL	PRFQELI	FEDPAF 575

Db	421	RVLERVLKKYDYDSSVVRKFRFREALIQISIPFLKKLAPCTCKSELPRQBELIFEDFAPR	480
Qy	576	ILVENVYEEVVLQTVWKMDILQAVKEAAVQRKENLYRDSWMHNSDNLHLHAEAPIDWG	635
Db	481	ILVENVYEEVVLQTVWKMDILQAVKEAAVQRKENLYRDSWMHNSDNLHLHAEAPIDWG	540
Qy	636	EYSNSGGGGSPSPSTPESATLSKRRRAKQVVVVQDEEVGLPPEASPEPPSPASPDGV	695
Db	541	EYSNSGGGGSPSPSTPESATLSKRRRAKQVVVVQDEEVGLPPEASPEPPSPASPDGV	600
Qy	696	TEIRGLLQAQGLRPESPPAPGLNCAAGESPOPKAAPEASSPPAPLQHLLPGKAVDLG	755
Db	601	TEIRGLLQAQGLRPESPPAPGLNCAAGESPOPKAAPEASSPPAPLQHLLPGKAVDLG	660
Qy	756	PKPSPDOETGEQVSSPSHPALHTTTEDSAGVQTF	791
Db	661	PKPSPDOETGEQVSSPSHPALHTTTEDSAGVQTF	696

RESULT 2
S26058
probable transforming protein (can) - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S26058
R: von Lindern, M.; Fornerod, M.; van Baal, S.; Jaegle, M.; de Wit, T.; Buijs, A.; Grosve-
Mol. Cell. Biol. 12, 1687-1697, 1992
A:Title: The translocation (6;9), associated with a specific subtype of acute myeloid le-
k-can mRNA.
A:Reference number: S26058; MUID:92195315; PMID:1549122
A:Accession: S26058
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2090 <VON>

Query Match	4.3%;	Score 176.5;	DB 2;	Length 2090;
Best Local Similarity	19.6%;	Pred. No. 0.098;		
Matches 148:	Conservative	126;	Mismatches 258;	Indels 225; Gaps 38;

154	QY	ROVPPRAVINSAGVKI	-----LTSVDQVLELIGNSLP-----GTTAKSGSAPILKCP	200
604	DB	QSAPEPSPFSASAKPAASGCLSHPTPLSAPPSSVPLKSLVSPSGRSAQSSSPVPSMV	663	
201	QY	TOFFPLIL	-----WHPYARHYFCMWTAEQDKQ-----AVLQDCIRHCNNGTPE--DS	247
664	DB	OKSPRIITPPAAKPGSPOAKSLQ	-----PAVAEKQCHQWKQSDPFWMAGIEEIAHFQKEULEELKA	722
248	QY	KVEGPAP	-----TDAIRMYRSK-----ELXGTWEMLCNGNEVOILSNLVMBELGPULK	295
723	DB	RTSKACFOVGTSSEMKNLRTESDDLHTFLLEIKETTESLHGDISLKTLLLEGFAGVEA	782	
296	QY	ALGPLRLKGPQERQORQWIIQISDAVY	-----HMVYE-----QAKARFEVLISKVQVQOPAMQAV	349
783	DB	RE	-----QNEENR-----DSGYLHLLYKRLPDKPSEAQLQIRLHQYVKFAVQ	826
350	QY	IRTDMDQIITSK	-----EHLASKIPAFILPKAEVCYRNHVQVPIPSILEALMWPTSQGTEVR	407
827	DB	DVNDVLDLEWDOHLEQKK	-----QRH-----LLVPE-----R	854
408	QY	DVFFKEVTDMLNVINEGGIDKLGBEYMEKLSRLAYHPLKMQSYKEMESLRLGLOQRD	467	
855	DB	ETLFNTLAN-NREIIN	-----QQRKRLNHL-----VDSLOQLRL	887
468	QY	VGSTVFKORAQIHMRQMDNAVYTFETILLHQELGCKPTEELCKSIQORLVERVLKKDY	527	
888	DB	YKQTSLWSLSAV	-----PSSSIHGFSDSL-----ESLCNAL-----LKTITISHTK	930
528	QY	DSSSVKRF	-----FREALLQISITPLLKLAPTKCSLSELP-----RFOELIFEDFAR	574
931	DB	SLPKVPAPKLSPMKQAKUR	-----NFLAKRKTTPVRSTAPASILSRSAFLSQRYEEDL-DEVSS	986
575	QY	FILVENTYBEVVLQTVMKD	-----ILQAVEAQVQRKHL-----YRDSVMWMSNDP	621

```

Db      987 TSSVSQSLSESDARTSKCDKDEAVQAPRHAPVVRTSIQPSLHLPAAFPKSHLVHGSSP 1046
      622 NL-----HLLAEGAPIDWGEYSNSGGSGSPSPSTPESA----- 655
Db      1047 GVMGTSVATSKAIPOGA--DSTMATKTVKHGAPSPSPHPI SAPQOLAAALRQMASQ 1104
      656 ----TLSEKRRARQVVVVQDEEV-----GLPFEASPEPPSPASPDGVT 697
Db      1105 APAVNTLTES--TLKNVPQVVNVQELKNPNPATPSTAMGSSVPYSTAKTTPHPVLTFVAANQ 1162
      698 IR-GLLAQGLRPESPPPA- GELLNGAPAGESPQKAA----PEASSPPASPLQLHLLPGKA 751
Db      1163 AKQGSLSLNLKPSGTPASQOLSGDGKASGTAKITAVTSTPSSAGQSKPFSFSPSGTG 1222
      752 VDLG--PPKPSDOETGEQVSPSPSHPALHTTTEDSAG 786
Db      1223 FNGIITPTPSSNFTAAOGATPSTKSSQPDAFSSGG 1259

RESULT 3
S15762
neurofilament triplet M protein - chicken
C:Species: Gallus gallus (Chicken)
C:Date: 20-Feb-1995 #sequence revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S15762; S08061; A27040
R:Zopf, D.; Dineva, B.; Betz, H.; Gundelfinger, E.D.
Nucleic Acids Res. 18, 521-529, 1990
A:Title: Isolation of the chicken middle-molecular weight neurofilament (NF-M) gene and
A:Reference number: S15762; PMID:90174973; PMID:2106668
A:Accession: S15762
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-858 <ZOP>
A:Cross-references: UNIPROT:P16053; UNIPARC:UPI000017137C; EMBL:X17102; NID:G63688; PID
R:Zopf, D.; Dineva, B.; Betz, H.; Gundelfinger, E.
submitted to the EMBL Data Library, November 1989
A:Reference number: S08061
A:Accession: S08061
A:Molecule type: DNA
A:Residues: 1-355,368-858 <ZOP>
A:Cross-references: UNIPARC:UPI00001774FC; EMBL:X17102
R:Zopf, D.; Hermans-Borgmeyer, I.; Gundelfinger, E.D.; Betz, H.
Genes Dev. 1, 699-708, 1987
A:Title: Identification of gene products expressed in the developing chick visual system
      327040. PMID:68112814; PMID:3123320

```

RESULT 3

S15762 neurofilament triplet M protein - chicken

C:Species: Gallus gallus (Chicken)

C:Date: 20-Feb-1995 #sequence revision 20-Feb-1995 #text_change 09-Jul-2004

C:Accession: S15762; S08061; A27040

R:Zopf, D.; Dineva, B.; Betz, H.; Gundelfinger, E.D.

Nucleic Acids Res. 19, 521-529, 1990

A:Title: Isolation of the chicken middle-molecular weight neurofilament (NF-M) gene and

A:Reference number: S15762; MUID:90174973; PMID:2106668

A:Accession: S15762

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-858 <ZOP>

A:Cross-references: UNIPROT:P16053; UNIPARC:UPI000017137C; EMBL:X17102; NID:G63688; PID

R:Zopf, D.; Dineva, B.; Betz, H.; Gundelfinger, E.

submitted to the EMBL Data Library, November 1989

A:Reference number: S08061

A:Accession: S08061

A:Molecule type: DNA

A:Residues: 1-355, 368-858 <ZOP>

A:Cross-references: UNIPARC:UPI00001774FC; EMBL:X17102

R:Zopf, D.; Hermans-Borgmeyer, I.; Gundelfinger, E.D.; Betz, H.

Genes Dev. 1, 699-708, 1987

A:Title: Identification of gene products expressed in the developing chick visual system

A:Reference number: A27040; MUID:88112814; PMID:3123320

A:Accession: A27040

A:Molecule type: mRNA

A:Residues: 350-546, 'R', 548-858 <ZOP>

A:Cross-references: UNIPARC:UPI000017137B; GB:X05558; NID:X05558; PIDN:CAA29073.1; PID:

C:Genetics:

A:introns: 355/3; 397/2

C:Superfamily: cytoskeletal

C:keywords: coiled coil

Query Match	3.9%	Score 162	DB 2	Length 858
Best Local Similarity	21.1%	Pred. No. 0.18		
Matches 136	Conservative 92	Mismatches 260	Indels 156	Caps 29
263	QSKELYGTWMLC--GNEVQIISLN-----VMEELGPELKAELPLPLGKGPQE	308		
QY				
Db	79	QSSLLNGAAELKILSRNKEQLQGNDRFGYTEKVHYLEQNKKEIAELA-ALRQKHAG	137	
QY	309	RQRQWIQISDAVHYMYEQAKRFEVLISKVQVQ-----	343	
Db	138	R---AQLGDA-YEQELRELRLGALBQNSHEKAQIQLDSBHI	192	
QY	344	--PAMQAVIRTDMDIITSKELHASKIRAFITLPKAEVCVVRNHVQYIPISILEALMVPQSO	401	
Db	193	ETETATIALRKEMEEASLWRAELDKVQS--LQDEAVAFILRGHNEBEVAELIAQLQ--ASH	248	
QY	402	GTEVRDVF-----PKEV-----TDMNLAVINEG--GIDKLGEYMEK-----	436	

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2006, 01:44:35 ; Search time 139.848 Seconds
(without alignments)
2586.080 Million cell updates/sec

Title: US-10-612-318-4
Perfect score: 4139
Sequence: 1 MEEVGRWGLKPLRGLRG.....SSHPALHTTTEDSAGVQTEF 791

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A. Genseqp2006s

- 1: Genseqp1980s:*
- 2: Genseqp1990s:*
- 3: Genseqp2000s:*
- 4: Genseqp2001s:*
- 5: Genseqp2002s:*
- 6: Genseqp2003as:*
- 7: Genseqp2003bs:*
- 8: Genseqp2004s:*
- 9: Genseqp2005s:*
- 10: Genseqp2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4139	100.0	791	ADL23468	Human STR
2	4114	99.4	821	ADL23466	Human STR
3	3824	92.4	733	AD055194	Protein #
4	3813	92.1	733	AA811456	Human Meg
5	3792	91.6	746	AA051209	Human ost
6	3792	91.6	746	ABM81075	Tumour-as
7	3662	88.5	703	ABP62933	Human pol
8	2027	49.0	396	AA095443	Human pro
9	1160	28.0	928	AA067451	Amino aci
10	1160	28.0	928	AA094034	Human pro
11	1160	28.0	928	ADY17313	PRO polyp
12	1160	28.0	928	ADY17311	PRO polyp
13	1160	28.0	928	AE04903	Cancer-as
14	1160	28.0	928	AE04901	Cancer-as
15	1160	28.0	928	AE07465	Human acc
16	1156	27.9	928	AD10972	Human the
17	989.5	23.9	246	AA064036	Human imm
18	838	20.2	726	AD082699	Human pro
19	827.5	20.0	725	AE04898	Cancer-as
20	658.5	15.9	383	AD04521	Human pro
21	658.5	15.9	383	AE04521	Human CDN
22	595	14.4	298	AA042724	Human ORF
23	542.5	13.1	703	ABR83120	Human BCN

24	522.5	12.6	620	7	ADM05460	Adm05460 Human pro
25	522.5	12.6	620	9	AEC88390	Aec88390 Human cDN
26	519.5	12.6	697	7	ABR83122	ABR83122 Human BCN
27	518.5	12.5	651	7	ABR83121	ABR83121 Human BCN
28	468	11.3	579	7	ADD67630	Human EY1
29	468	11.3	579	7	ABR83123	ABR83123 Human BCN
30	329	7.9	247	4	AA090354	Human imm
31	244	5.9	177	2	AA021849	Human sig
32	177	4.3	2130	9	AD07779	Cyclin-de
33	176.5	4.3	2090	10	AE058534	Human nuc
34	176.5	4.3	2090	10	AE058534	Human nuc
35	176.5	4.3	2127	7	ADP09542	Human pol
36	175.5	4.2	2058	4	AA097070	Human pro
37	166	4.0	31	9	ADV13344	Human pro
38	162.5	3.9	873	8	AD055126	Protein #
39	161.5	3.9	1259	8	ADN23114	Bacterial
40	161	3.9	1278	8	AD182558	Human mod
41	159.5	3.9	1248	2	AA013464	Human dia
42	158.5	3.8	1315	2	AA076734	Human mbi
43	156.5	3.8	1082	7	AD014368	Human int
44	156.5	3.8	1088	8	ADH61284	INTSIG pr
45	156	3.8	533	4	AA095397	Human pro

ALIGNMENTS

RESULT 1
ADL23468
ID ADL23468 standard; protein; 791 AA.
XX
AC ADL23468;
XX
DT 01-JUL-2004 (first entry)
XX
DE Human STR50 short variant polypeptide.
XX
KW Human; STR50; short variant; neurotoxic stress;
KW neurodegenerative disease; stroke; Parkinson's disease;
KW Alzheimer's disease; Huntington's disease; tumour; cancer;
KW cerebroprotective; vasotonic; antiparkinsonian; neuroprotective;
KW nontropic; anticonvulsant.
XX
OS Homo sapiens.
XX
PN US2004068098-A1.
XX
PD 08-APR-2004.
XX
PF 01-JUL-2003; 2003US-00612318.
XX
PP 01-JUL-2002; 2002US-0393251P.
XX
PR (KACH/) KACHALSKY S G.
XX
PA (PAER/) FAERMAN A.
XX
PE (PELO/) PEL-OR Y.
XX
PI Kachalsky SG, Faerman A, Pel-Or Y;
XX
DR WPI; 2004-304622/28.
XX
DR N-PSDB; ADL23467.
XX
PT New STR50 polynucleotides and polypeptides, useful in diagnosing and
PT treating neurodegenerative diseases, e.g. stroke, Parkinson's disease,
PT Alzheimer's disease or Huntington's disease and tumors.
XX
PS Claim 15; SEQ ID NO 4; 38pp; English.
XX
CC The invention relates to a purified polynucleotide encoding the human
CC STR50 polypeptide. The purified polypeptide has the biological activity
CC to modulate neurotoxic stress. The polynucleotide and the polypeptide, and
CC compositions of the invention are useful in diagnosing and treating
CC neurodegenerative diseases, e.g. stroke, Parkinson's disease, Alzheimer's

